

Proteomics and Application of Mass Spectrometry

Kessiri Kongmanas, Ph.D.

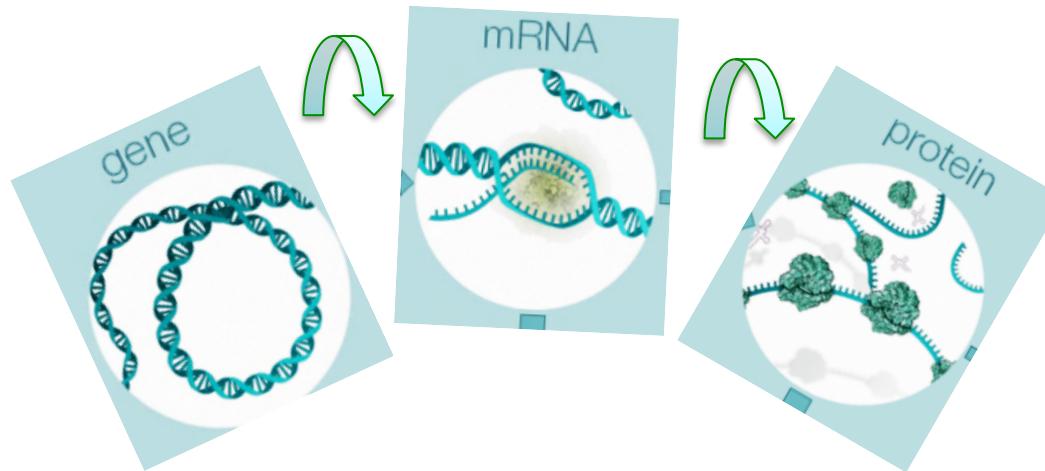
SIRE 503 Medical Bioinformatics
25 September 2018

Outline

- **Definition (Proteome & Proteomics)**
- **Progress of proteomic technology**
- **Mass Spectrometry (MS)**
- **Principles of MS-based proteomics**
- **Quantitative proteomics**
- **Application of proteomics**

The proteome is far more complex than the genome.

Proteome: The protein expression profile of a cell, an organism, or a tissue under exactly defined conditions.

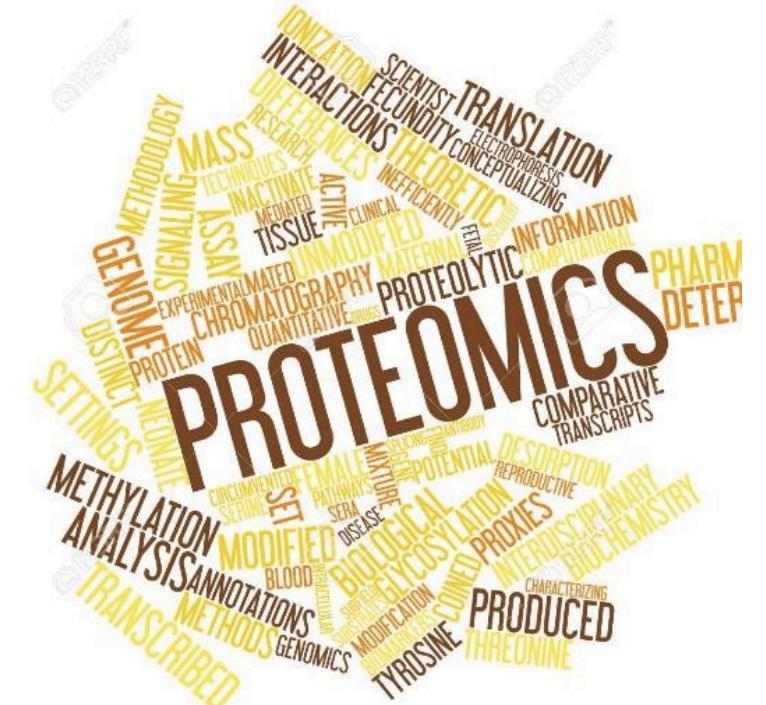


The release of the Human Genome Sequence data in 2004 →
Humans are considered to have ~19,559 genes encoding proteins

Genes (~19,559 genes) → Alternative RNA splicing + Post translational modification → **≥ 1 million Proteins**

Proteomics

- The term “proteomics” was coined from merging “protein” and “genomics” in the 1990s.
 - Proteomics is the study of proteome.
 - The study of proteins that are associated with a disease by means of their altered levels of expression and/or post-translational modification between control and disease states.



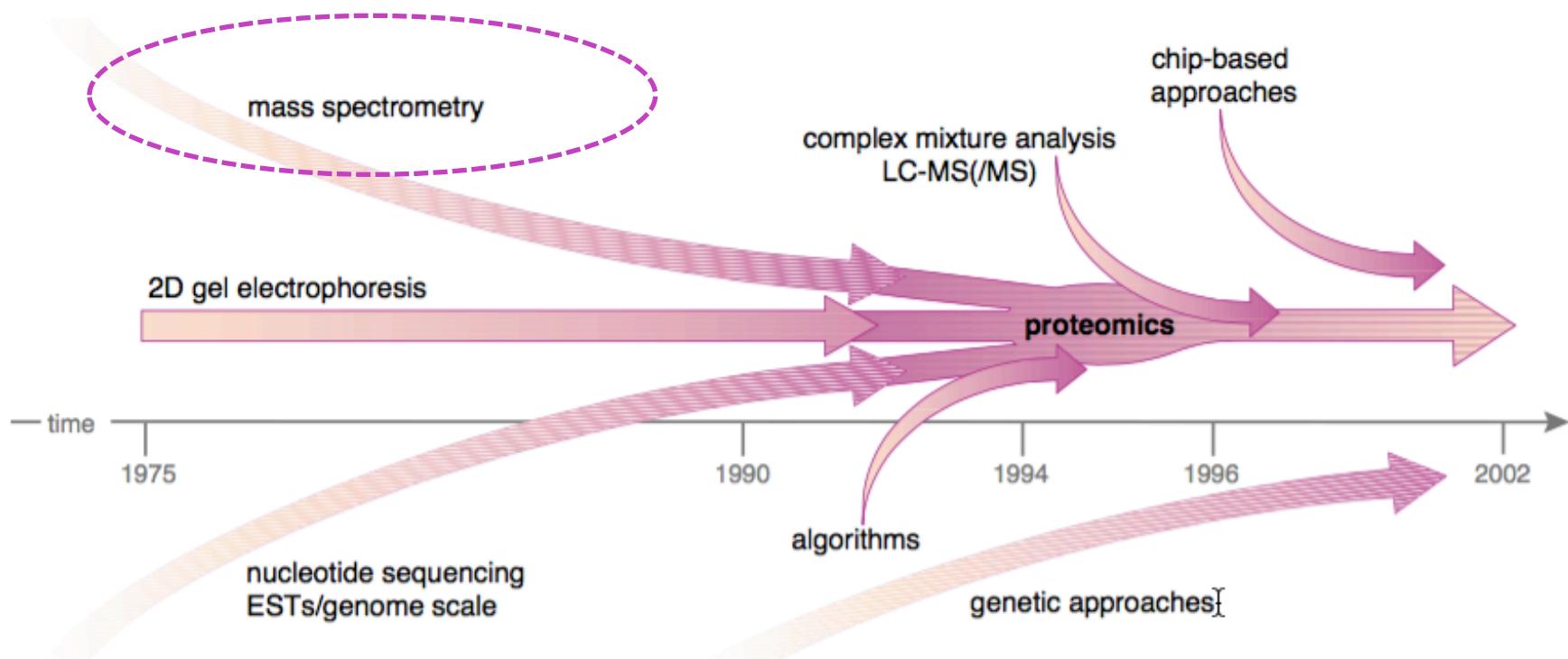
Data from proteomics



The range of proteins produced by cell/tissue VS

**The initiation or progression
of a disease state and the
effect of therapy.**

Timeline of the convergence of different technologies and resources into a proteomic process

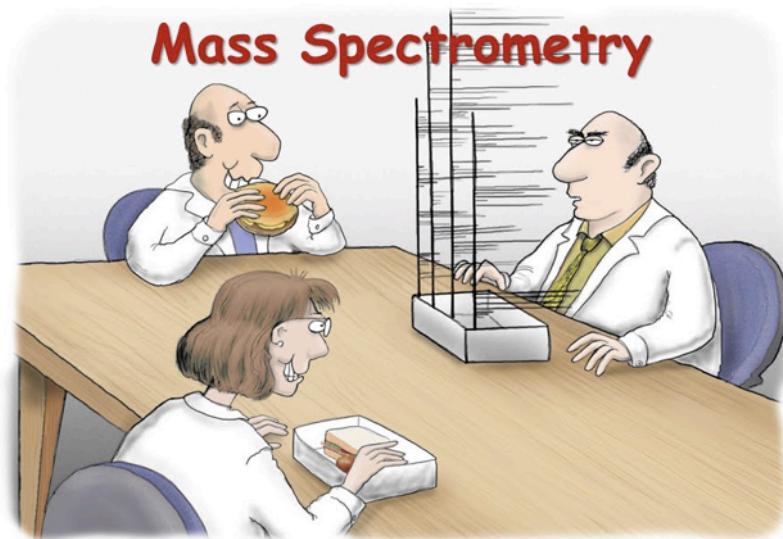


Nat Genet. 2003 Mar;33 Suppl:311-23.

Mass Spectrometry (MS)

An analytical technique that ionizes chemical species and sorts the ions based on their mass-to-charge ratio (m/z).

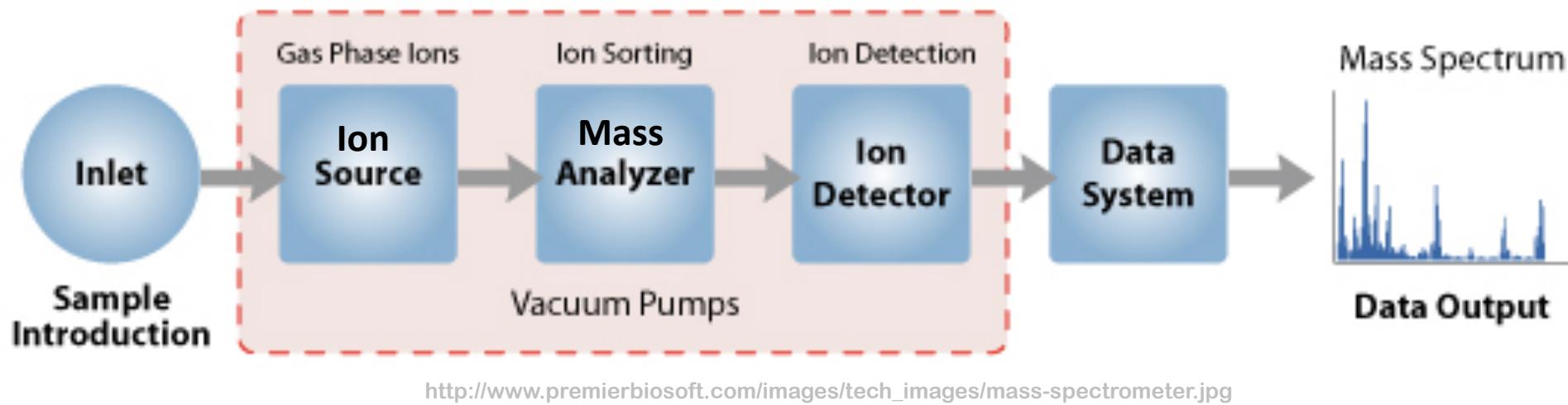
From Wikipedia, the free encyclopedia



Courtesy www.lab-initio.com

<https://i.ytimg.com/vi/ruT-xBsnxok/maxresdefault.jpg>

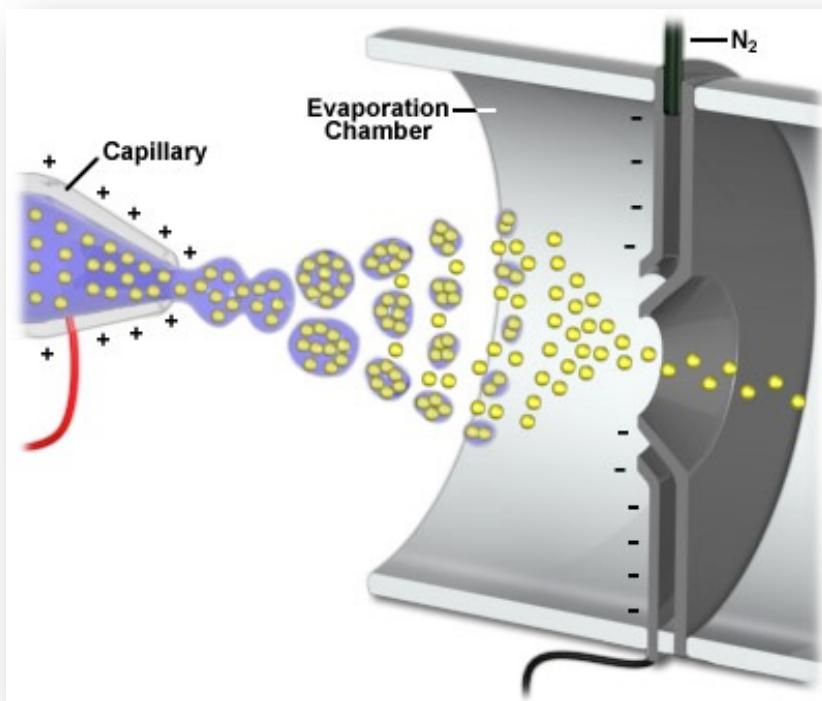
Basic components in a mass spectrometer



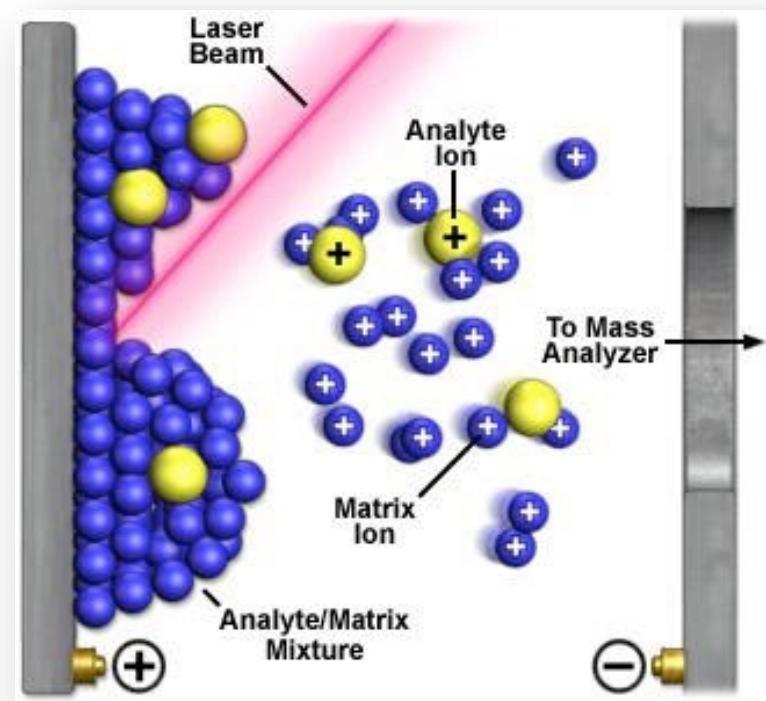
- The molecules need to be converted into ions first because electric and magnetic fields are used to manipulate the ions, uncharged species (molecules) do not respond. Ions are manipulated by their charge, separated by their mass to charge ratio (m/z).
- Vacuum is needed because the gas phase ion should pass through the analyzer without fragmenting or having significant trajectory alteration by collisions with any molecules or other ions.

Ion Source

Two major ionization methods used in proteomics



ESI



MALDI

Old ESI-MS (1960s)



Single quadrupole mass spectrometer used for John Fenn Nobel Prize winning work on electrospray ionization

Ancient MS (before 1900)



New ESI-MS (2000s)

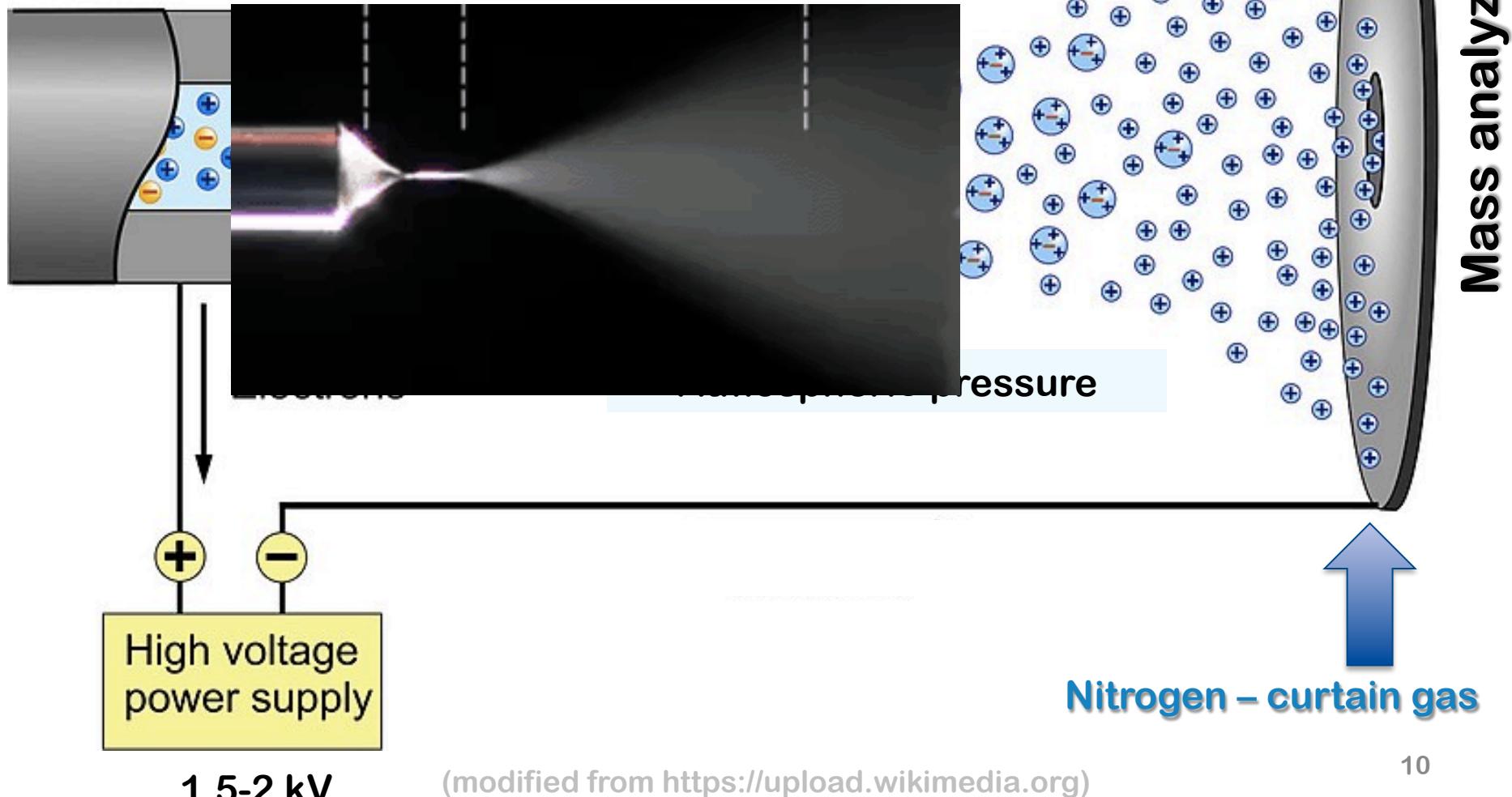


Thermo Scientific - LTQ Orbitrap XL™ ETD
Hybrid Ion Trap-Orbitrap Mass Spectrometer

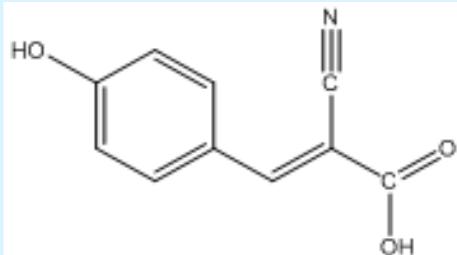
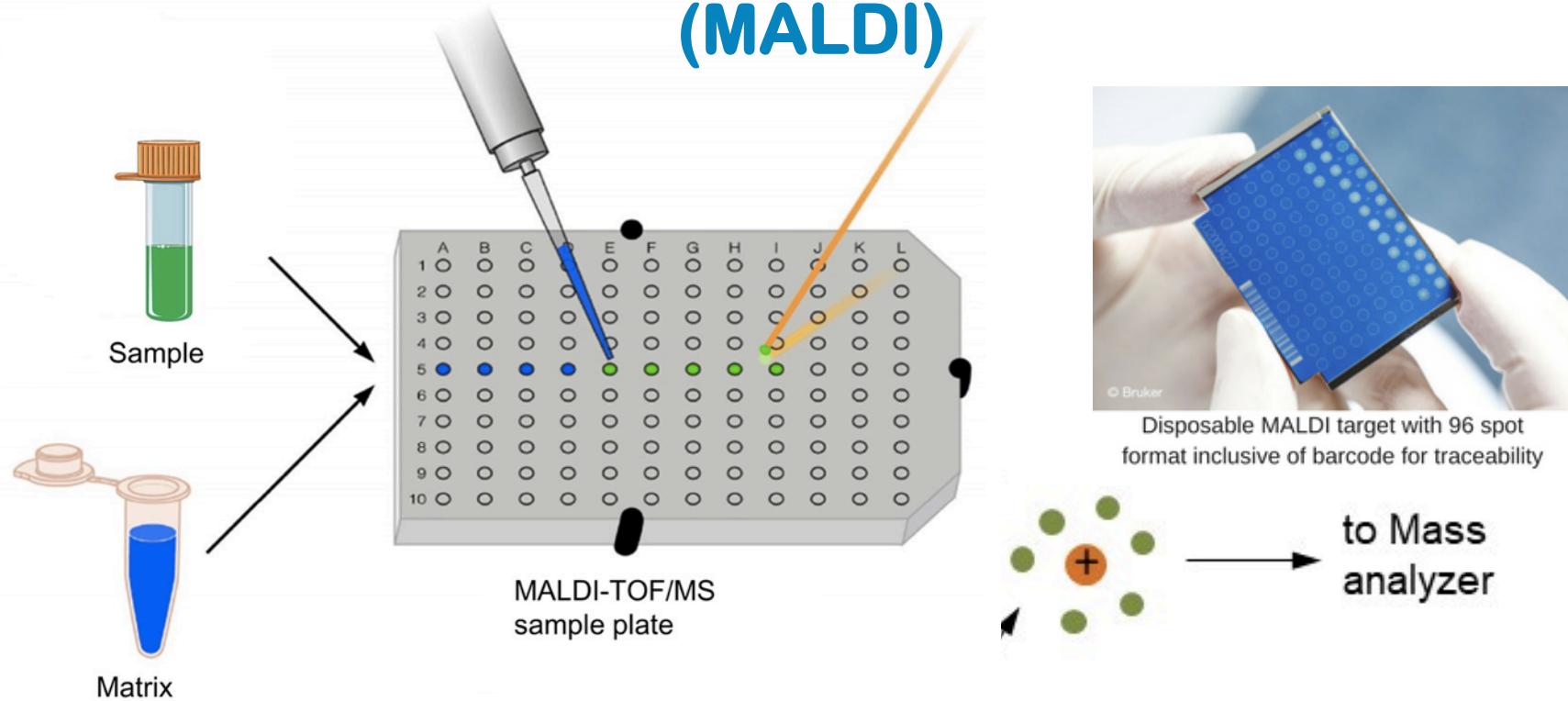
Electrospray ionization (ESI)

Droplet size ↓ → Charge density ↑ → Ions ejected

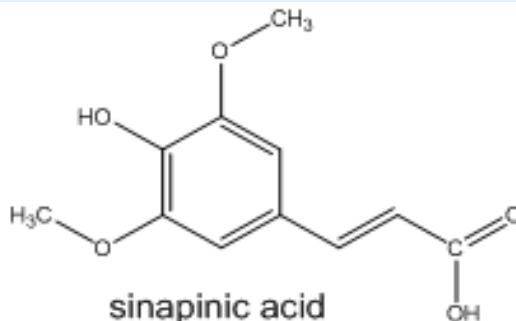
Sample in aqueous buffer sprayed from the Taylor cone



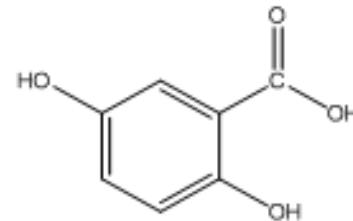
Matrix-Assisted Laser Desorption Ionization (MALDI)



α -cyano-4-hydroxycinnamic acid
CCA $C_{10}H_7NO_3$
peptides and small molecules



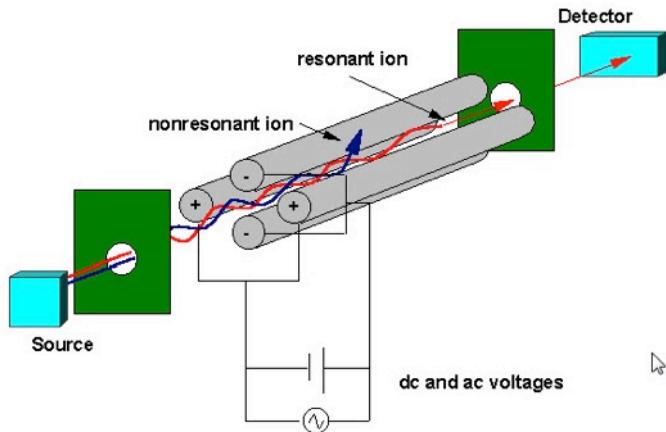
sinapinic acid
SA $C_{11}H_{12}O_5$
proteins



2,5 dihydroxybenzoic acid
DHB $C_7H_6O_4$
oligosaccharides

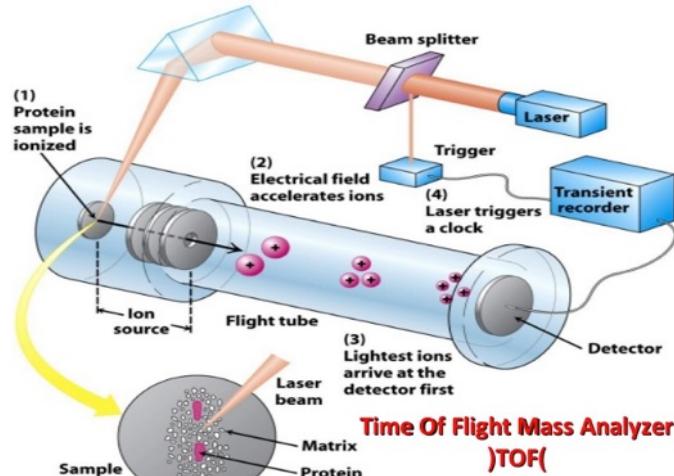
Different Types of Mass Analyzer

Quadrupole (Q)



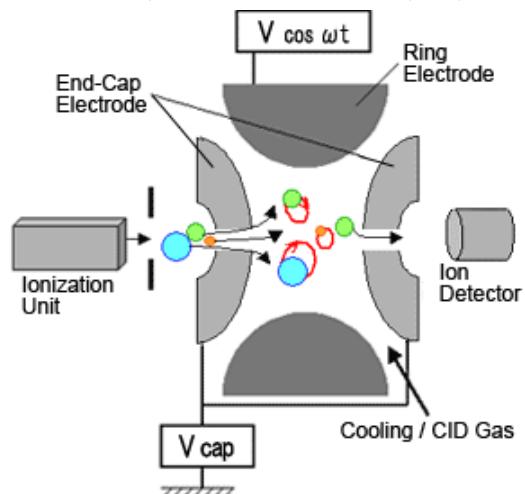
<http://www.cif.iastate.edu/sites/default/files/uploads/MS/Quadrupole%20pic.jpg>

Time-of-Flight (ToF)



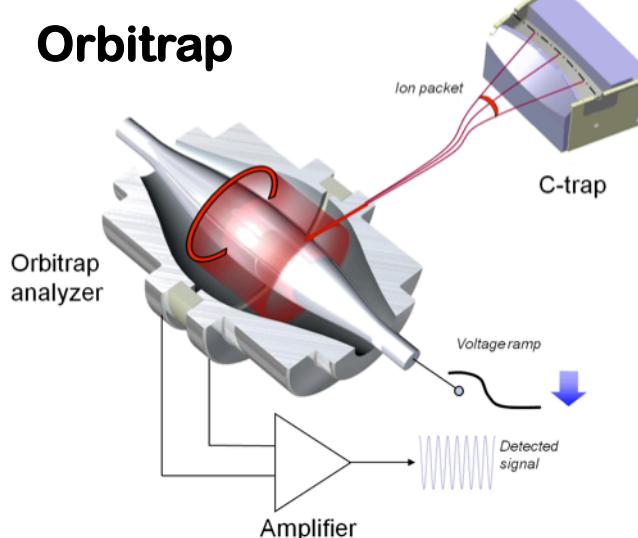
<http://image.slidesharecdn.com/massspectrometrybasicprinciples-140309102217-phapp0295/mass-spectrometry-basic-principles-20-638.jpg?cb=1394360615>

Quadrupole ion trap (QIT)



https://encrypted-tbn1.gstatic.com/images?q=tbn:ANd9GcSOet0Qogy0S1GqOgw5yF0IRW_Ri-ANMrVQCkDarS0DGnDHF4K

Orbitrap



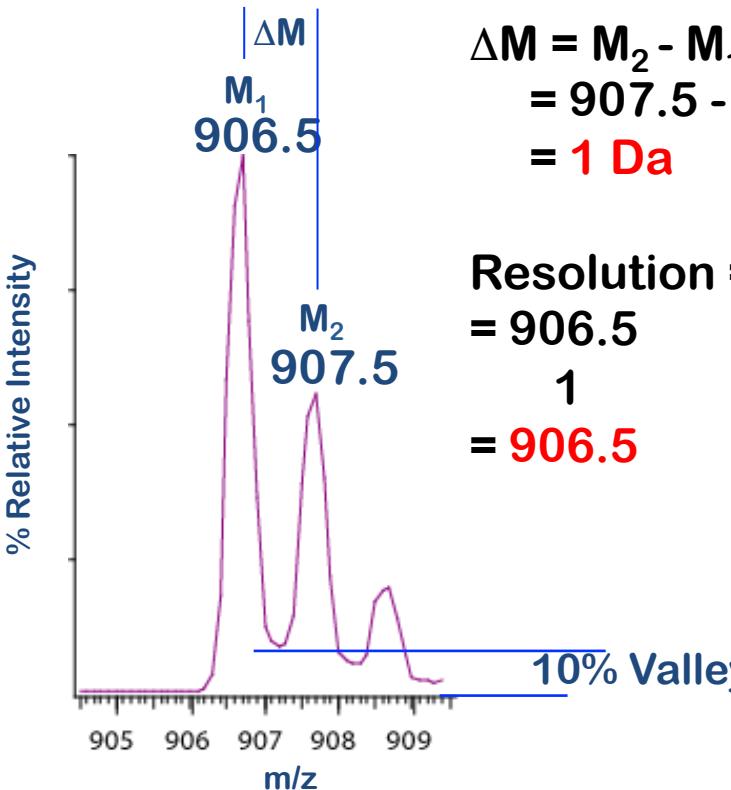
Mass Analyzers

- Quadrupole (Q): low resolution, inexpensive, fast scanning.
- Quadrupole ion traps (QIT): low resolution, inexpensive, fast scanning can also be used in the tandem mode with some limitations.
- Linear ion traps (LIT): inexpensive, fast scanning, medium resolution, used as a stand-alone analyzer or as a trapping cell in a tandem instrument.
- Time-of-flight (TOF): high resolution, fast scanning, constantly being improved.
- Ion cyclotron resonance cell – Fourier Transform (FTMS): very high resolution, very expensive.
- Orbitrap: new kid on the block, very high resolution, expensive (but less so than FTMS).

Resolution

Resolution: Ability of the mass spectrometer to distinguish between ions of similar m/z.

$$\text{Resolution} = \frac{M}{\Delta M}$$

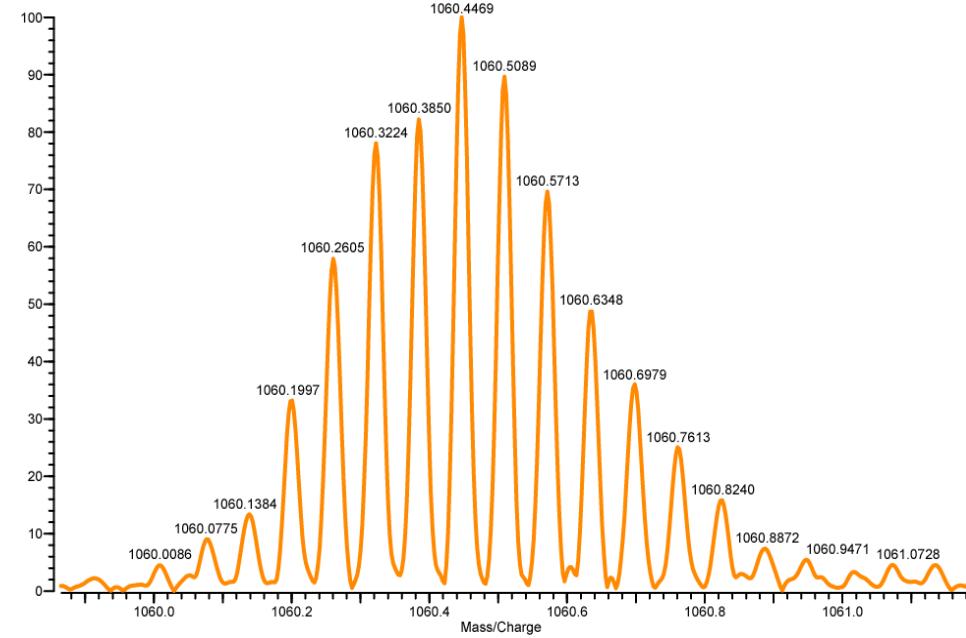


$$\begin{aligned}\Delta M &= M_2 - M_1 \\ &= 907.5 - 906.5 \\ &= 1 \text{ Da}\end{aligned}$$

$$\begin{aligned}\text{Resolution} &= \frac{M}{\Delta M} \\ &= 906.5 \\ &\quad 1 \\ &= 906.5\end{aligned}$$

$$\Delta M = 1060.4469 - 1060.3850 = 0.0619$$

$$\frac{M}{\Delta M} = \frac{1060.4469}{0.0619} = 17,131$$



Data Collected on ESI-Quad Sciex API III+

Data Collected on ESI-FTMS IonSpec 7T

Mass accuracy & Calibration

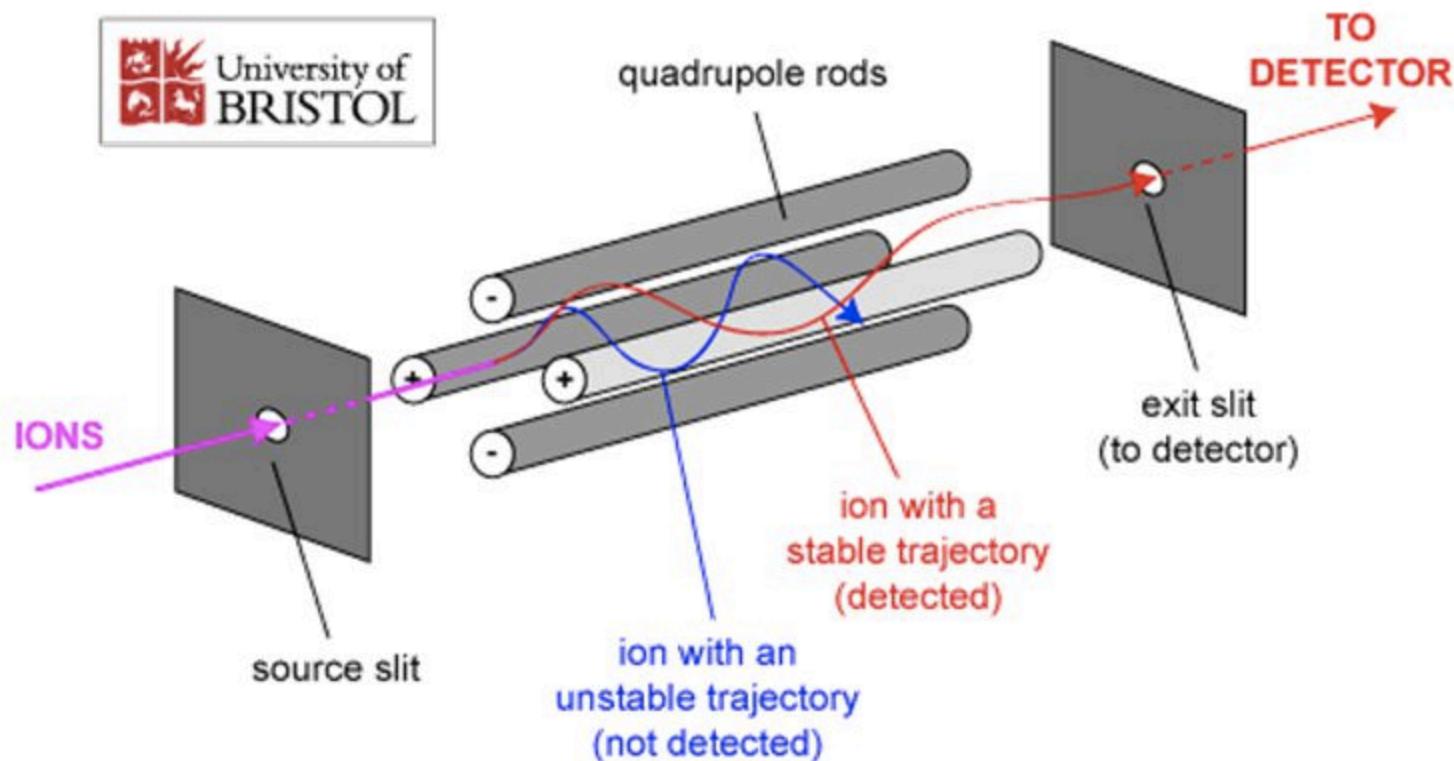
- **Accuracy of measurement:** The degree of conformity of a measured quantity to its actual true value.
- Mass measurement errors are expressed at parts-per-million [ppm] or mDa (Best measurement today is 1 ppm.)
 - $\text{ppm} = \frac{\text{observed m/z} - \text{calculated m/z}}{\text{calculated m/z}} \times 10^6$
- **Calibration (to get the accurate mass)**
 - External Calibration
 - Instrument calibrated and then sample of interest introduced
 - Good for most routine work
 - Internal Calibration
 - A compound of known mass is introduced along with sample of interest
 - In every spectrum the exact (known) m/z value of the calibrant is assigned
 - The m/z values of all other ions in each spectrum is then corrected
 - Used when maximum accuracy required

Power of exact mass

Elemental composition possibilities: C [0-500], H [0-1000], N [0-6], O [0-6], P [0-3], and S [0-4]

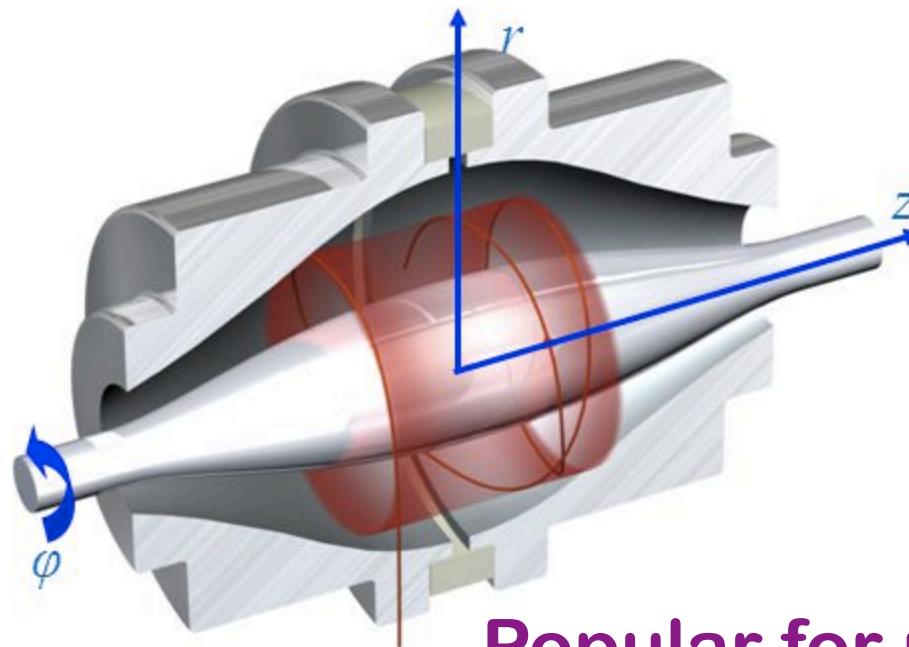
	If your instrument has this capability:			
m/z [Experimental]	< 2ppm	< 5ppm	< 10ppm	< 20ppm
100.0258	$\text{CH}_2\text{N}_5\text{O}$	$\text{CH}_2\text{N}_5\text{O}$	$\text{CH}_2\text{N}_5\text{O}$	$\text{CH}_2\text{N}_5\text{O}$ H_6NO_5 $\text{C}_3\text{H}_4\text{N}_2\text{O}_2$
275.2689	$\text{C}_{13}\text{H}_{33}\text{N}_5\text{O}$	$\text{C}_{13}\text{H}_{33}\text{N}_5\text{O}$ $\text{C}_{15}\text{H}_{35}\text{N}_2\text{O}_2$	$\text{C}_{13}\text{H}_{33}\text{N}_5\text{O}$ $\text{C}_{15}\text{H}_{35}\text{N}_2\text{O}_2$	5 possibilities
773.5624	$\text{C}_{48}\text{H}_{76}\text{N}_3\text{O}_3\text{P}$ and 29 others!!	53 possibilities	100 possibilities	Too many to count!

Quadrupole



- Can be thought as a mass filter
- DC and AC fields applied that stabilize a trajectory for ions in a desired mass range, undesired ions are ejected
- Quadrupole operate with a continuous flow of ions
- Low resolution (± 1 Da)

The Orbitrap Mass Analyzer

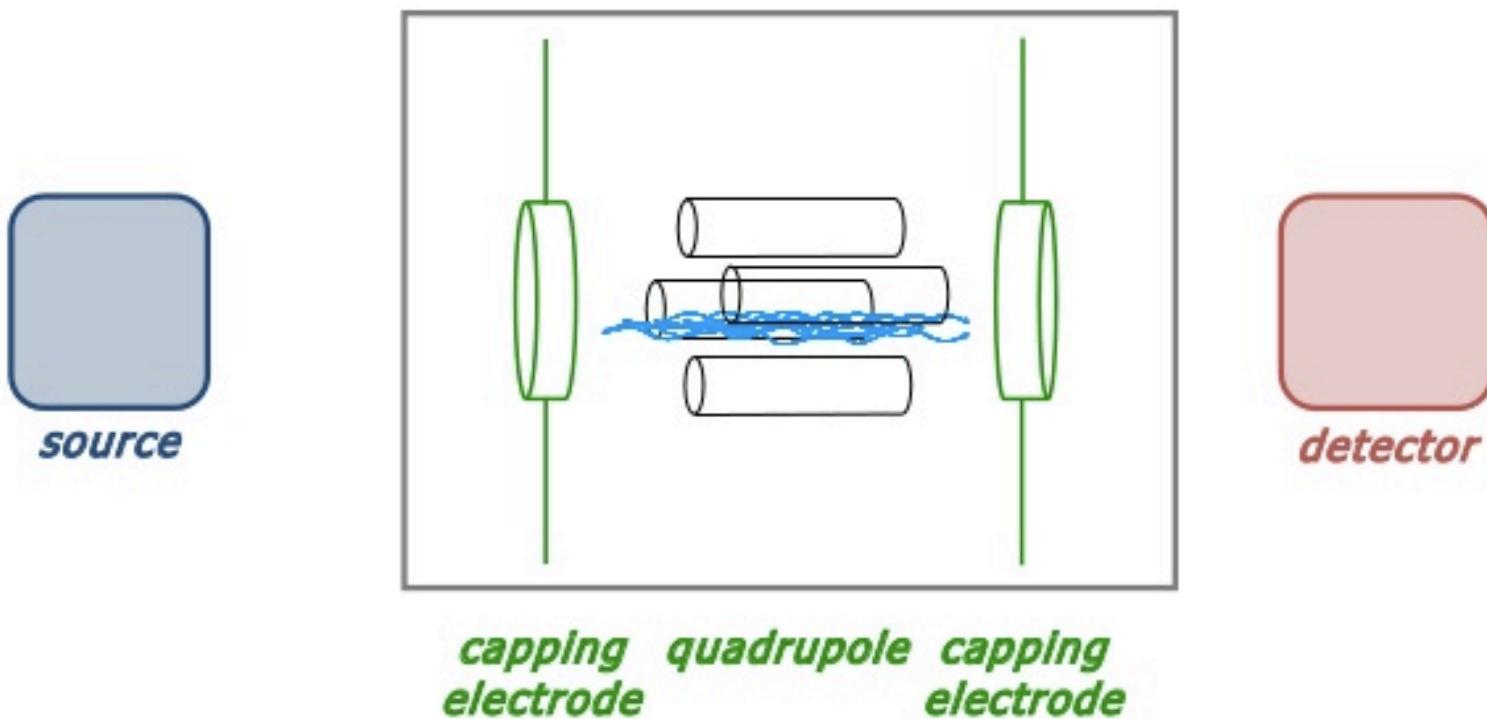


$$\omega_z = \sqrt{\frac{k}{m/q}}$$

Popular for proteomic work

- The first Orbitrap was introduced in 2005
- The Orbitrap is a Fourier transform mass spectrometer
- Ions oscillate at a frequency proportional to m/z
- Image current detection produces a “transient” that is converted to a mass spectrum via a Fourier transform.
- High resolution, mass accuracy, and throughput

ESI quadrupole linear ion trap



Linear ion traps store ions in a fixed linear trajectory rather than in a 'blob'.

The major advantage is the increased capacity for trapping ions. This allows for a broader dynamic range and better quantitation performance. The linear ion trap is these days often encountered as the workhorse mass spectrometer in many labs, and can be extended with the highly accurate orbitrap or FT-ICR mass analyzers/detectors.

Time-of-Flight Mass Spectrometry

To determine m/z values

A packet of ions is accelerated by a known potential and the flight times of the ions are measured over a known distance.

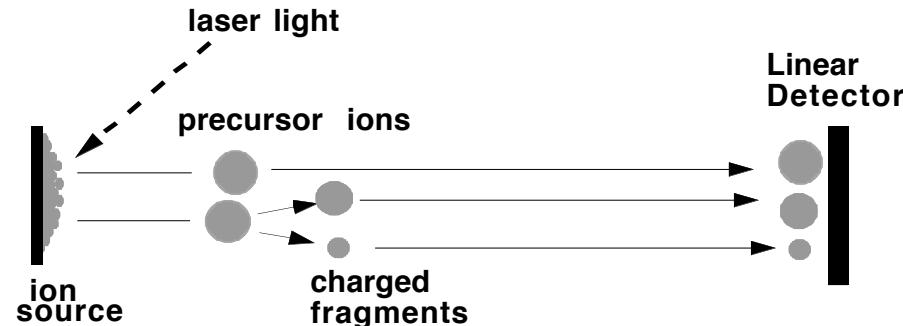
Q: What are V, e, and z?

Key Performance Notes

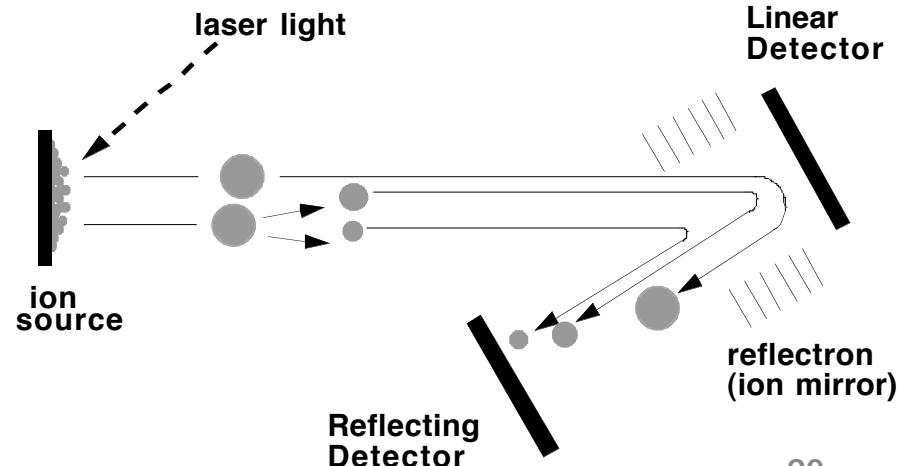
- Based on dispersion in time
- Measures all m/z simultaneously, implying potentially high duty cycle
 - “Unlimited” mass range
 - DC electric fields
 - Small footprint
 - Relatively inexpensive

Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry

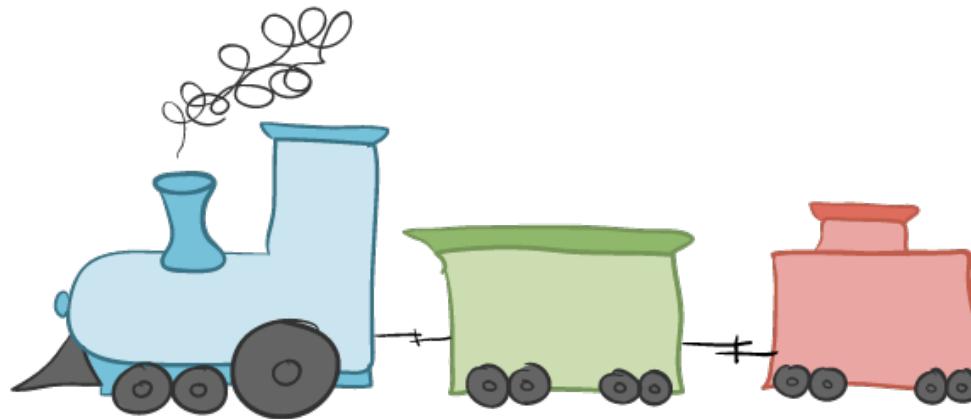
Linear Time-of-Flight



Reflecting Time-of-Flight

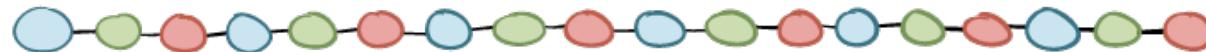


MS-BASED PROTEOMIC ANALYSIS



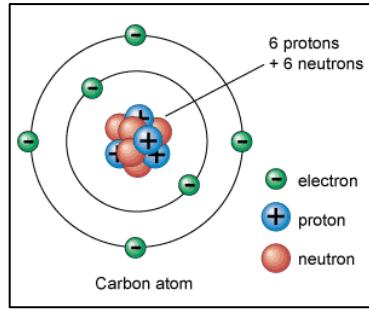
<https://ka-perseus-images.s3.amazonaws.com/9faccce4d1ba3e31815036ba8b97439e58857656.svg>

Amino acids link together like the cars of a train



Periodic Table of the Elements

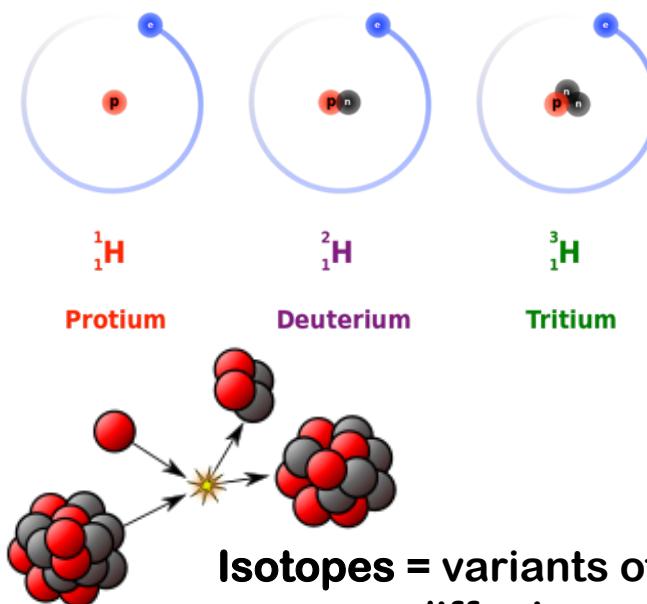
1 H Hydrogen 1.01	2 He Helium 4.00
3 Li Lithium 6.94	4 Be Beryllium 9.01
11 Na Sodium 22.99	12 Mg Magnesium 24.31
19 K Potassium 39.10	20 Ca Calcium 40.08
21 Sc Scandium 44.96	22 Ti Titanium 47.88
23 V Vanadium 50.94	24 Cr Chromium 51.99
25 Mn Manganese 54.94	26 Fe Iron 55.93
27 Co Cobalt 58.93	28 Ni Nickel 58.69
29 Cu Copper 63.55	30 Zn Zinc 65.39
31 Ga Gallium 69.73	32 Ge Germanium 72.61
33 As Arsenic 74.92	34 Se Selenium 78.09
35 Br Bromine 79.90	36 Kr Krypton 84.80
37 Rb Rubidium 84.49	38 Sr Strontium 87.62
39 Y Yttrium 88.91	40 Zr Zirconium 91.22
41 Nb Niobium 92.91	42 Mo Molybdenum 95.94
43 Tc Technetium 98.91	44 Ru Ruthenium 101.07
45 Rh Rhodium 102.91	46 Pd Palladium 106.42
47 Ag Silver 107.87	48 Cd Cadmium 112.41
49 In Indium 114.82	50 Sn Tin 118.71
51 Sb Antimony 121.76	52 Te Tellurium 127.6
53 I Iodine 126.90	54 Xe Xenon 131.29
55 Cs Cesium 132.91	56 Ba Barium 137.33
57-71 Lanthanides	72 Hf Hafnium 178.49
73 Ta Tantalum 180.95	74 W Tungsten 183.85
75 Re Rhenium 186.21	76 Os Osmium 190.23
77 Ir Iridium 192.22	78 Pt Platinum 195.08
79 Au Gold 196.97	80 Hg Mercury 200.59
81 Tl Thallium 204.38	82 Pb Lead 207.20
83 Bi Bismuth 208.98	84 Po Polonium [208.98]
85 At Astatine 209.98	86 Rn Radon 222.02
87 Fr Francium 223.02	88 Ra Radium 226.03
89-103 Actinides	104 Rf Rutherfordium [261]
105 Db Dubnium [262]	106 Sg Seaborgium [266]
107 Bh Bohrium [264]	108 Hs Hassium [269]
109 Mt Meitnerium [268]	110 Ds Darmstadtium [269]
111 Rg Roentgenium [272]	112 Cn Copernicium [277]
113 Uut Ununtrium unknown	114 Fl Flerovium [289]
115 Uup Ununpentium unknown	116 Lv Livermorium [298]
117 Uus Ununseptium unknown	118 Uuo Ununoctium unknown
57 La Lanthanum 138.91	58 Ce Cerium 140.12
59 Pr Praseodymium 140.91	60 Nd Neodymium 144.24
61 Pm Promethium 144.91	62 Sm Samarium 150.36
63 Eu Europium 151.97	64 Gd Gadolinium 157.25
65 Tb Terbium 158.93	66 Dy Dysprosium 162.50
67 Ho Holmium 164.93	68 Er Erbium 167.26
69 Tm Thulium 168.93	70 Yb Ytterbium 173.04
71 Lu Lutetium 174.97	
89 Ac Actinium 227.03	90 Th Thorium 232.04
91 Pa Protactinium 231.04	92 U Uranium 238.03
93 Np Neptunium 237.05	94 Pu Plutonium 244.05
95 Am Americium 243.06	96 Cm Curium 247.07
97 Bk Berkelium 247.07	98 Cf Californium 251.08
99 Es Einsteinium [254]	100 Fm Fermium 257.10
101 Md Mendelevium 258.10	102 No Nobelium 259.10
103 Lr Lawrencium [262]	



Alkali Metal Alkaline Earth Transition Metal Basic Metal Semimetal Nonmetal Halogen Noble Gas Lanthanide Actinide

Mass of elements

- All elements have a unique mass
- Mass comes in three flavors
 - Integer
 - Monoisotopic
 - Average



Isotopes = variants of a particular chemical element which differ in neutron number.

Element	Integer Mass	Monoisotopic Mass	Isotopic abundance	Average Mass
C	12	12	${}^{12}\text{C}$ 100%	12.011
			${}^{13}\text{C}$ 1.08%	
H	1	1.00782504	${}^1\text{H}$ 100%	1.00794
			${}^2\text{H}$ 0.016%	
N	14	14.00307	${}^{14}\text{N}$ 100%	14.00674
			${}^{15}\text{N}$ 0.4%	
O	16	15.99491464	${}^{16}\text{O}$ 100%	15.9994
			${}^{17}\text{O}$ 0.04%	
			${}^{18}\text{O}$ 0.2%	
P	31	30.9737634	${}^{31}\text{P}$ 100%	30.973763
S	32	31.9720718	${}^{32}\text{S}$ 100%	32.069
			${}^{33}\text{S}$ 0.8%	
			${}^{34}\text{S}$ 4.4%	

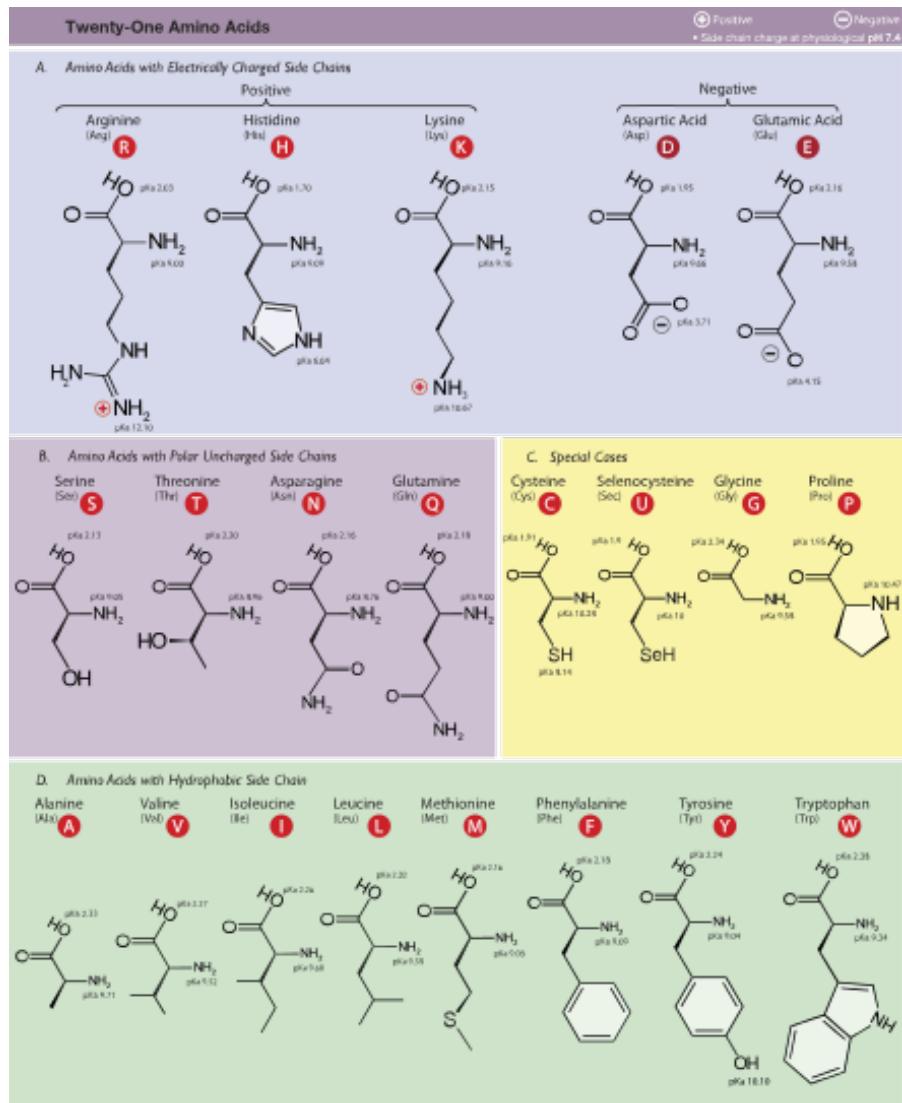
Amino acids have unique masses.

Name	Symbol	Integer Average	Monoisotopic
Alanine	Ala	71	71.0788
Arginine	Arg	156	156.1876
Asparagine	Asn	114	114.1039
Aspartic acid	Asp	115	115.0886
Cysteine	Cys	102	103.1448
*Glutamine	Gln	128	128.1308
Glutamic acid	Glu	129	129.1155
Glycine	Gly	57	57.0520
Histidine	His	137	137.1412
**Isoleucine	Ile	113	113.1595
**Leucine	Leu	113	113.1595
*Lysine	Lys	128	128.1742
Methionine	Met	131	131.1986
Phenylalanine	Phe	147	147.1766
Proline	Pro	97	97.1167
Serine	Ser	87	87.0782
Threonine	Thr	101	101.1051
Tryptophan	Trp	186	186.2133
Tyrosine	Try	163	163.1760
Valine	Val	99	99.1326
			99.06841

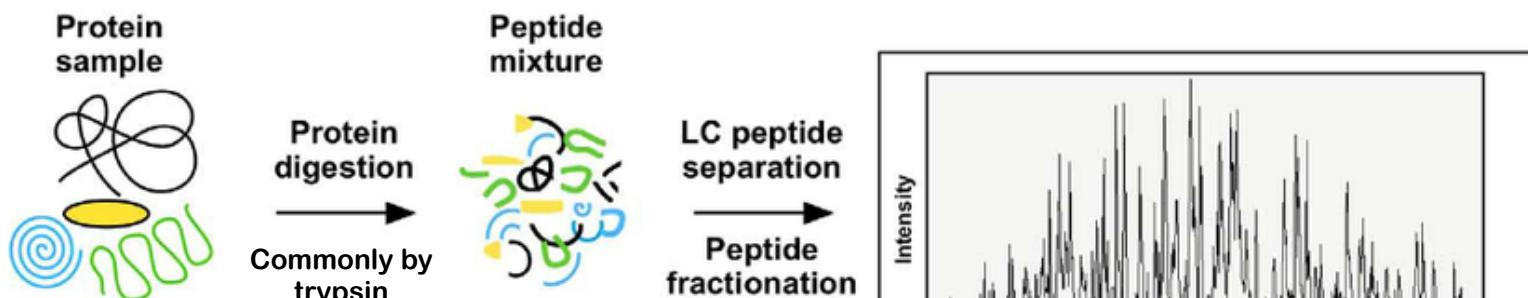
Exceptions:

**Ile and Leu have same elemental composition.

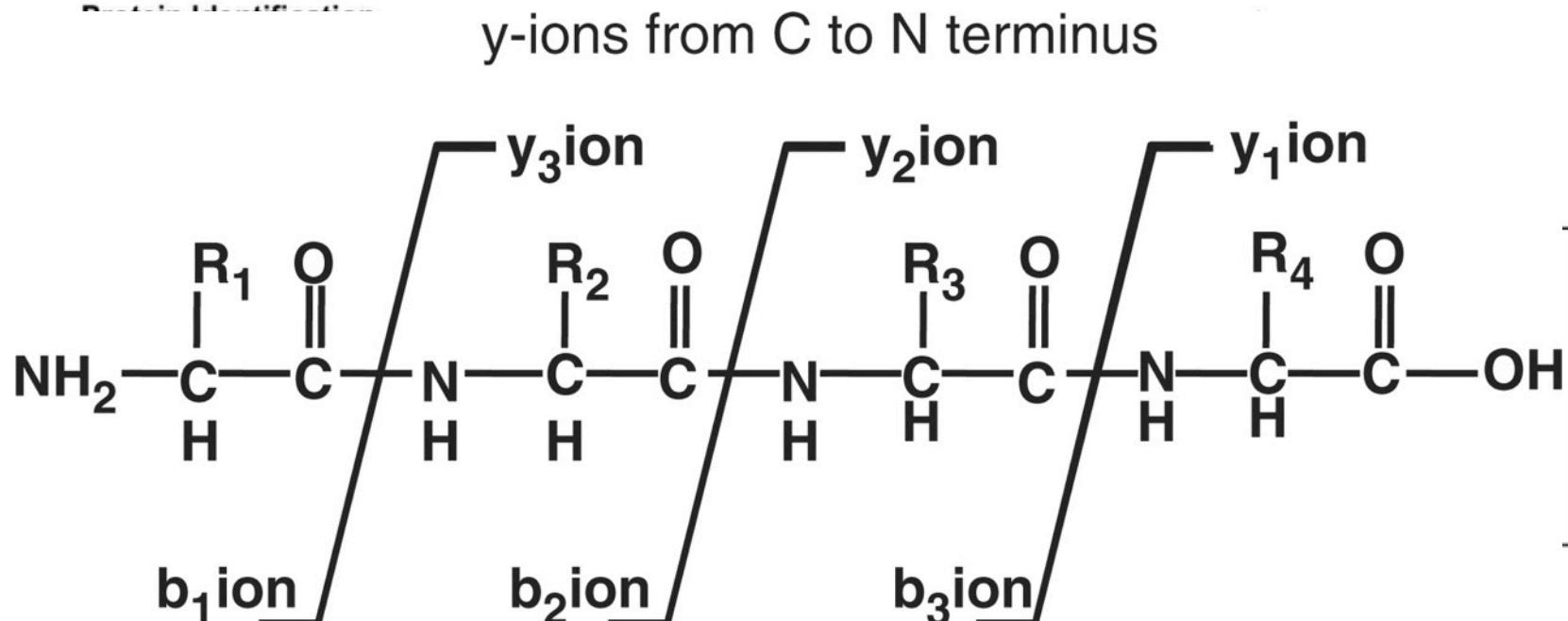
*Lys and Gln have similar masses.



https://upload.wikimedia.org/wikipedia/commons/thumb/a/a9/Amino_Acids.svg/440px-Amino_Acids.svg

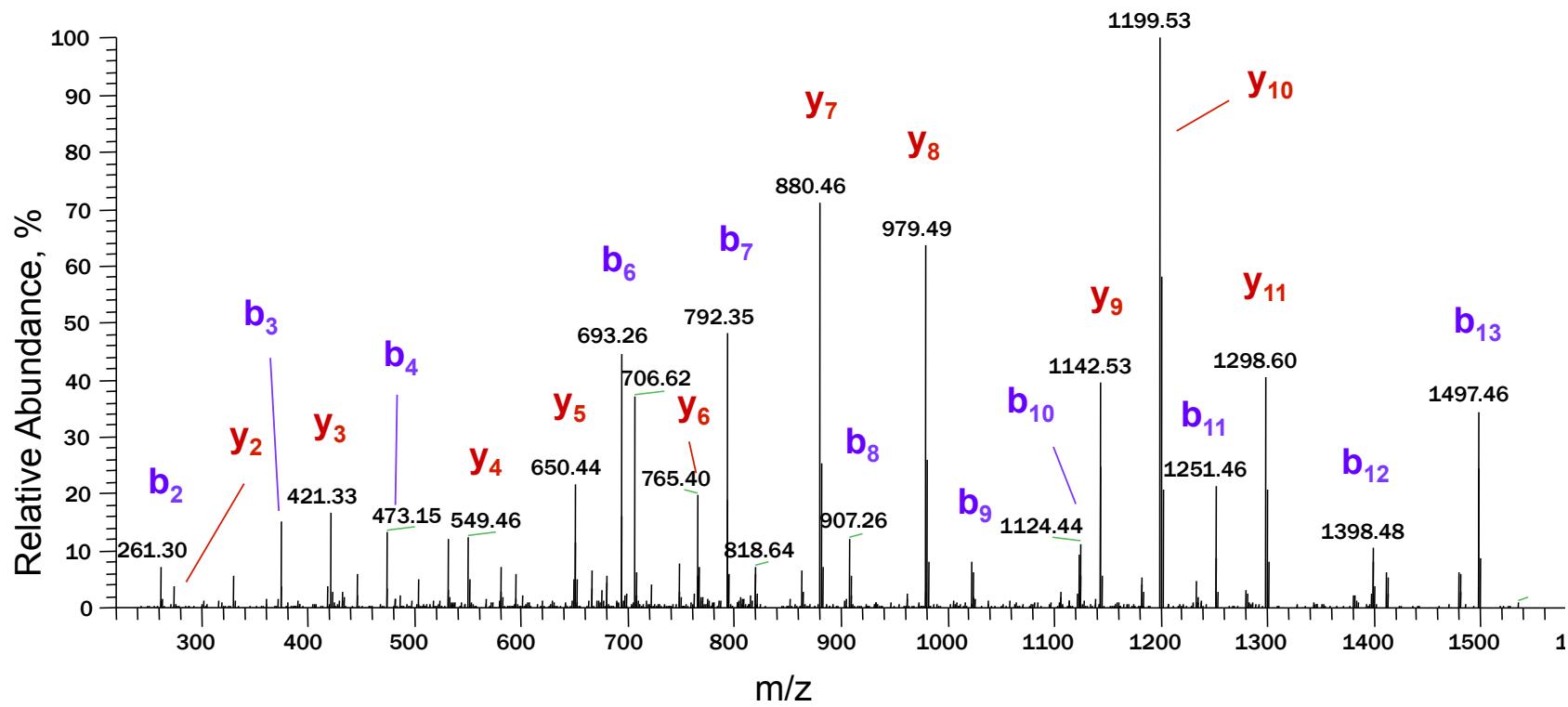


Collisionally activated dissociation (CAD) of peptides

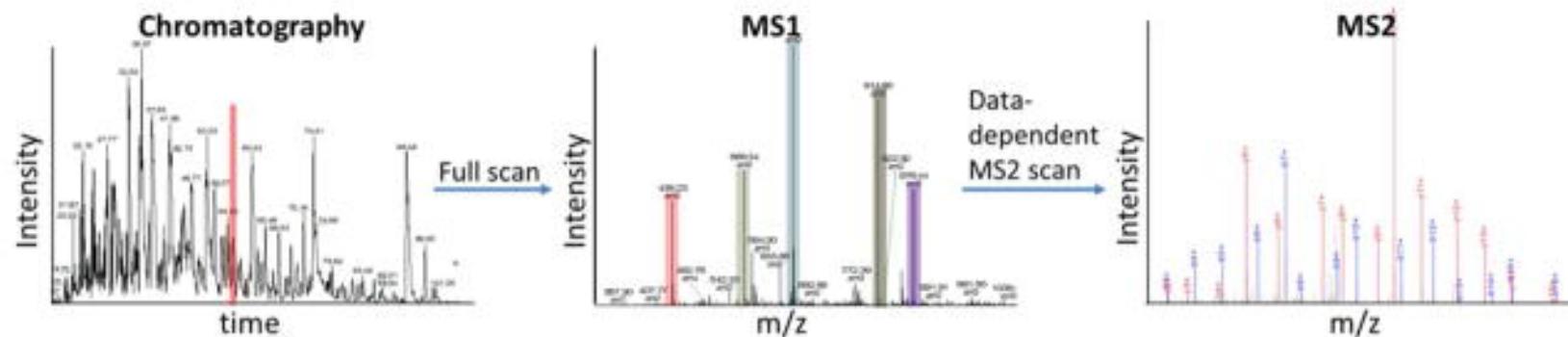


Example electrospray MS/MS spectrum of a peptide

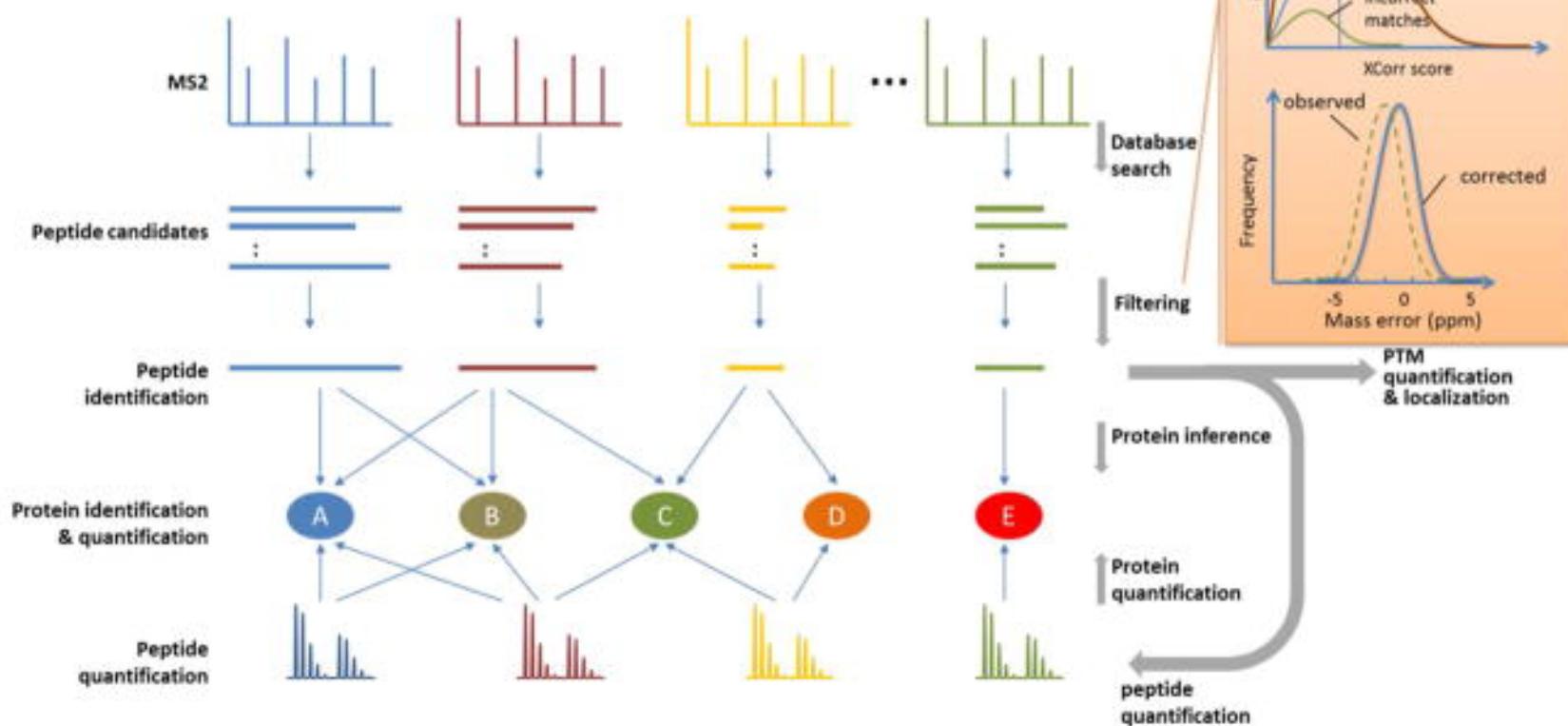
11 10 9 8 7 6 5 4 3 2 y ions
F-I-I-V-G-Y-V-D-D-T-Q-F-V-R
b ions 2 3 4 6 7 8 9 10 11 12 13



a: Data acquisition

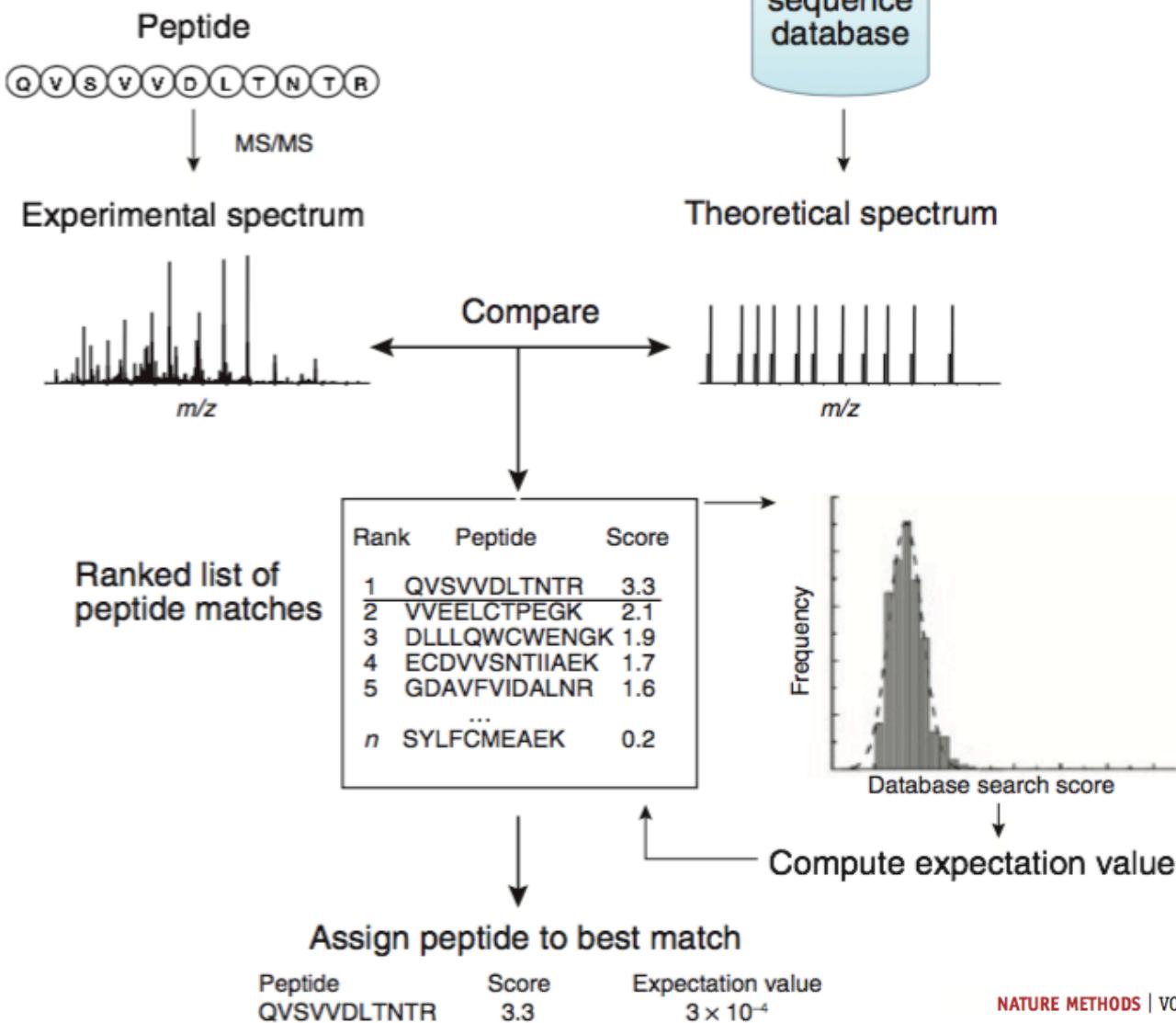


b: Data analysis



Protein databases

- NCBI
- UniProt



<https://www.ncbi.nlm.nih.gov/protein/?term=actin>https://www.ncbi.nlm.nih.gov/protein/NP_001284915.1https://www.ncbi.nlm.nih.gov/protein/NP_001284915.1

```

<region_name="PTZ00281"
</note="actin; Provisional"
</db_xref="CDD:173506"
Region      6..376
</region_name="ACTIN"
</note="Actin; smart00268"
</db_xref="CDD:214592"
Site        order(12..15,17,19,138,155..158)
</site_type="other"
</note="nucleotide binding site [chemical binding]"
</db_xref="CDD:212657"
CDS        1..376
</gene="Act5C"
</locus_tag="Dmel_CG4027"
</gene_synonym="A; A4V404_DROME; Ac5C; act; Act; ACT; act
5C; Act-5C; ACT1_DROME; Act5; act5C; Act5c; ACT5C; actin;
Actin; Actin/BAP47; actin5C; Actin5C; anon-EST:fe2D2;
Bap47; BAP47; beta-actin; beta-actin/Bap47; CG4027;
chrX:5748184..5748304; cyt5C; Dmel\CG4027; l(1)G0009;
l(1)G0010; l(1)G0025; l(1)G0079; l(1)G0117; l(1)G0177;
l(1)G0245; l(1)G0330; l(1)G0420; l(1)G0486; M32055; T11"
</coded_by="NM_001297986.1:162..1292"
</db_xref="FLYBASE:FBpp0311818"
</db_xref="FLYBASE:FBgn0000042"
</db_xref="GeneID:31521"
```

ORIGIN

```

1 mcdeevaalv vdngsgmcka gfagddapra vfpsivgrpr hggvmvgmqq kdsyvgdeaq
61 skrgiltlyk piehgivtnw ddmekiwhht fynelrvape ehpvliteap lnpanrekm
121 tqimfetfnt pamvyaiqav lslyasgrtt givldsgdgv shtvpiyegy alphailrid
181 lagrdltdyl mkiltegrys fttaereiv rdikeklcyy aldfqemataassssleks
241 yelpdgqvit ignerfrccpe alfqpssflqm eachgihetty nsimkcdvdi rkdllyantvl
301 sggttmypgi adrmqkeita lapstmkiki iapperkysv wiggsilasl stfqgmwisk
361 qeydesgpsi vrkrkf
//
```

Sequenceⁱ

Sequence statusⁱ: Complete.

Sequence processingⁱ: The displayed sequence is further processed into a mature form.

P68133-1 [UniParc]

[FASTA](#)

[Add to basket](#)

[« Hide](#)

10	20	30	40	50
MCDEDETTAL VCDNGSGLVK AGFAGDDAPR AVFPSIVGRP RHQGVMVGMG				
60	70	80	90	100
QKD SYVGDEA QSKRGIL TLK YPIEHGIITN WDDMEKIWHH TFYNELRVAP				
110	120	130	140	150
EEHPTLLTEA PLNP KANREK MTQIMFETFN VPAMYVAIQA VLSLYASGR				
160	170	180	190	200
T GIVL DSGDG VTHN VPIYEG YALPHAIMRL DLAGRDLTDY LMKIL TERGY				
210	220	230	240	250
SFVT TAEREI VRDIKEKLCY VALDFENEMA TAASSSSLEK SYELPDGQVI				
260	270	280	290	300
TIGNERFRCP ETLFQPSFIG MESAGIHETT YNSIMKCDID IRKDLYANNV				
310	320	330	340	350
MSG GTTM MPG IADRM QKEIT ALAPSTM KIK IIAPPERKYS VWIGGSILAS				
360	370			
LSTF QQMWIT KQEY DEAGPS IVHRKCF				

Length: 377

Mass (Da): 42,051

Last modified: July 21, 1986 - v1

Checksum: ⁱ DF2A3A046346A179

BLAST

GO

Natural variant

Database search programs and their websites

Program

Mascot

Masslynx

MS-Tag/MS-Seq

PeptideSearch

PepFrag

ProBID

SEQUEST

SpectrumMill

X!Tandem

Web site

<http://www.matrixscience.com/>

<http://www.waters.com>

<http://prospector.ucsf.edu/>

<http://www.narrador.embl-heidelberg.de/GroupPages/Homepage.html>

<http://prowl.rockefeller.edu/PROWL>

<http://projects.systemsbiology.net/probid>

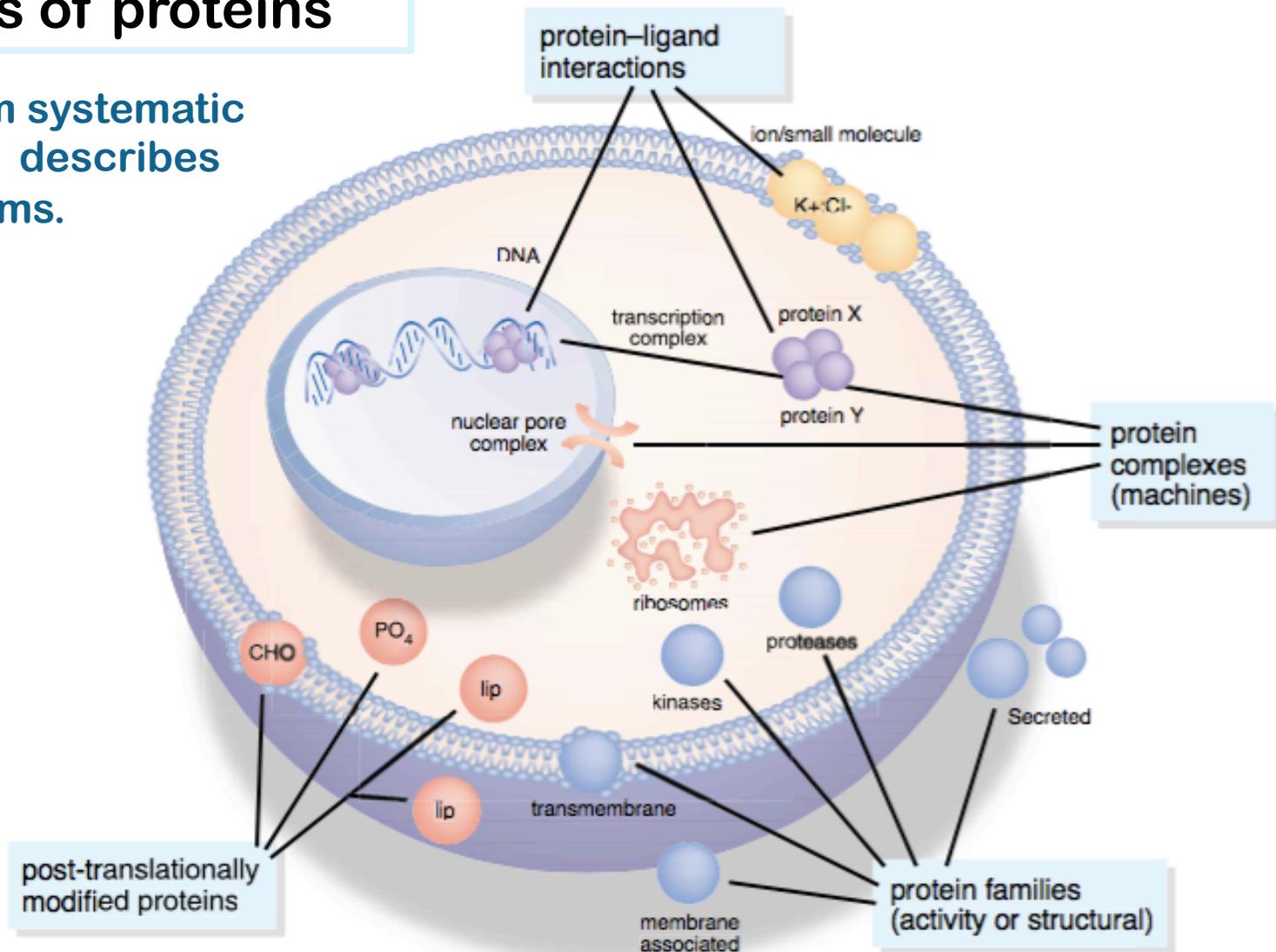
<http://fields.scripps.edu> or <http://www.thermo.com>

<http://www.chem.agilent.com/>

<http://www.thegpm.org/TANDEM/index.html>

Multidimensional properties of proteins

Information from systematic proteomics that describes biological systems.

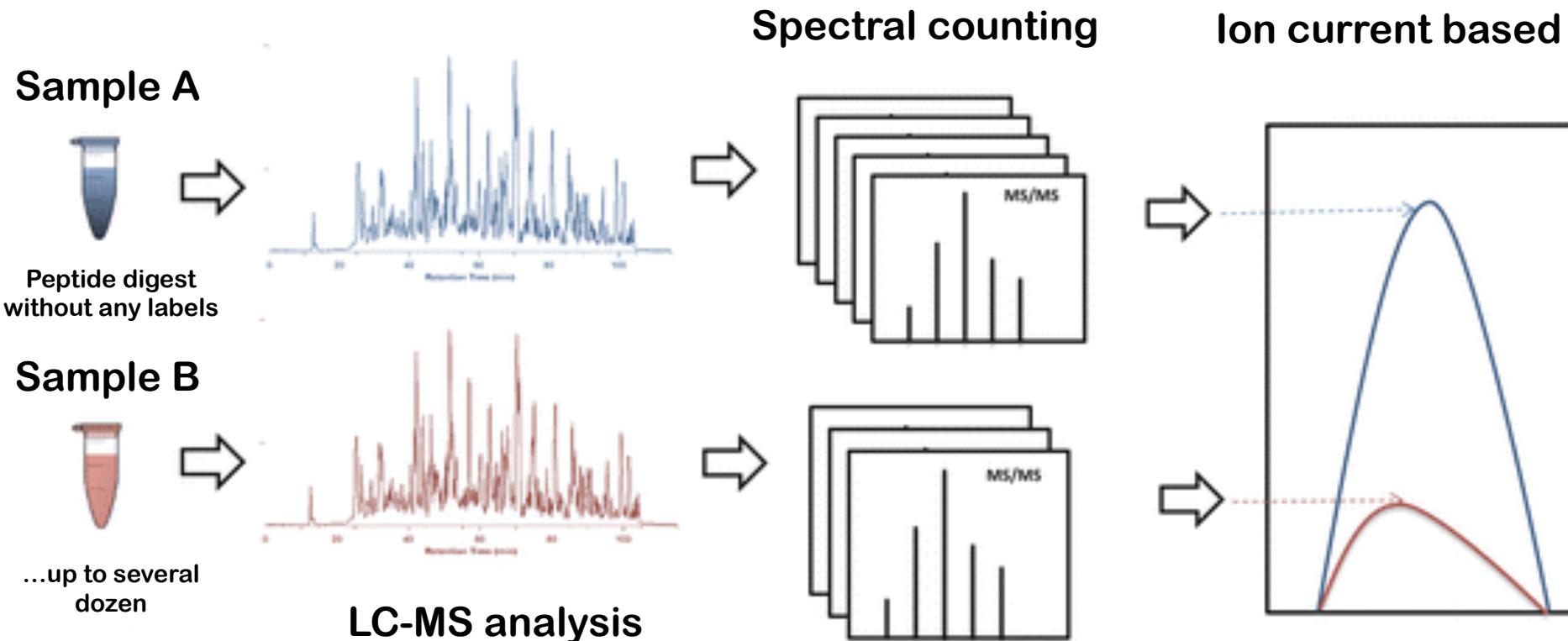
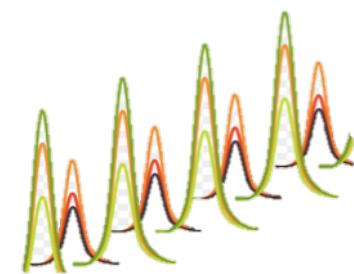


Protein dimension annotation resources (selected examples)

- Encyclopedia of Proteome Dynamics
(<https://peptracker.com/accounts/login/epd/>)
- Human Protein Atlas
(<https://www.proteinatlas.org/humancell>)
- Human Proteome Map
(<http://humanproteomemap.org/>)
- ProteomicsDB
(<https://www.proteomicsdb.org/>)
- The proteins across organisms database
(<https://pax-db.org/>)
- Protein interaction
(<https://string-db.org/>)

Quantitative Proteomics

Label-free quantitative proteomics

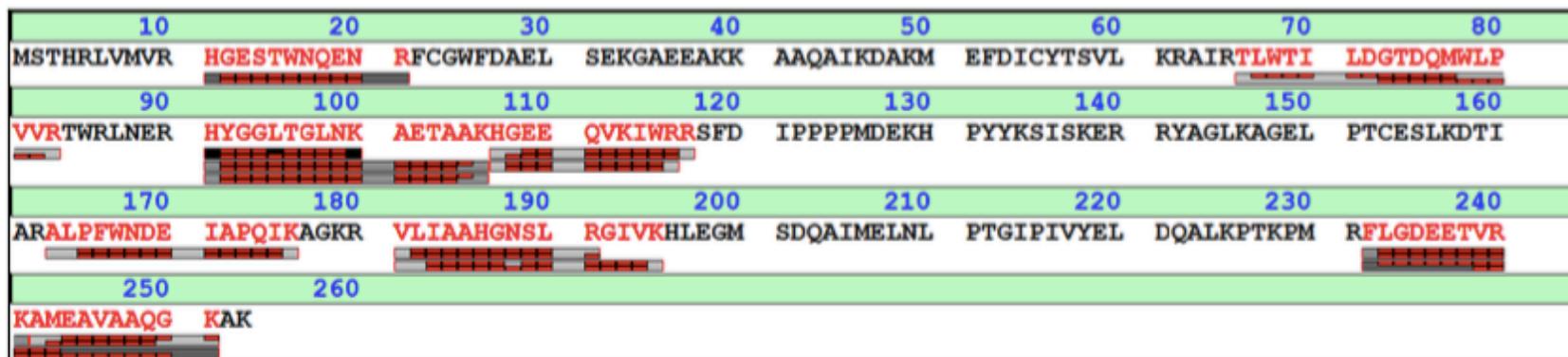


https://media.springernature.com/lw785/springer-static/image/chp%3A10.1007%2F978-3-319-31828-8_11/MediaObjects/331890_1_En_11_Fig3_HTML.gif

An example of label-free proteomic results using BioTools (Bruker's software)

Protein 3: Phosphoglycerate mutase 2 OS=Bos taurus GN=PGAM2 PE=2 SV=1

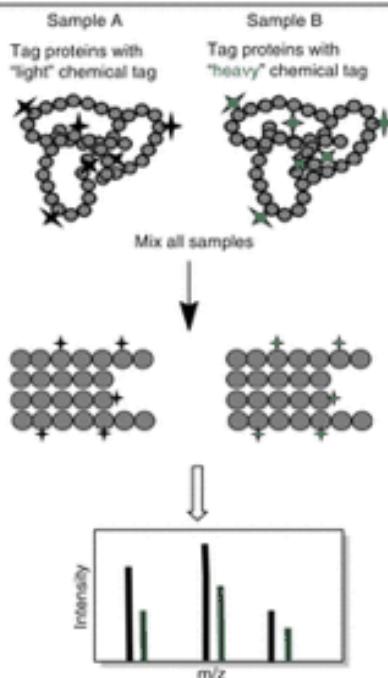
Accession: PGAM2_BOVIN **Score:** 834.0
Database: SwissProt(SwissProt_57.15.fasta) **MW [kDa]:** 28.7
Database Date: 2011-08-04 **pl:** 9.6
Modification(s): Deamidated **Sequence Coverage [%]:** 41.5
No. of unique Peptides: 13



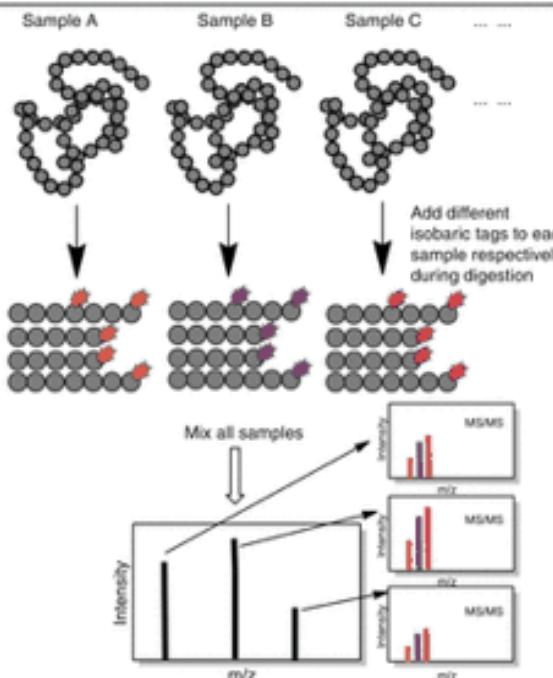
Cmpd.	No. of Cmpds.	m/z meas.	Δ m/z [ppm]	z	Rt [min]	Score	P	Range	Sequence	Modification
124353	1	679.0587	-351.72	2	124353.	77.0	0	11-21	R.HGESTWNQEN.R	
127148	1	715.2582	-168.13	3	127148.	56.1	0	66-83	R.TLWTILDQDMWLPVVR.T	
124660	1	530.1636	-224.64	2	124660.	52.6	0	91-100	R.HYGGLTGLNK.A	
36038	1	815.6387	-358.25	2	36038.0	99.9	1	91-106	R.HYGGLTGLNKAETAAK.H	
36049	1	817.2587	1023.59	2	36049.0	89.7	1	91-106	R.HYGGLTGLNKAETAAK.H	Deamidated: 9
36062	1	641.3636	39.04	2	36062.0	53.4	1	107-116	K.HGEEQVKIWR.R	
35818	1	719.2386	-209.23	2	35818.0	38.3	2	107-117	K.HGEEQVKIWR.R.S	
126263	2	821.0587	-458.58	2	126263.	92.0	0	163-176	R.ALPFWNDEIAPQIK.A	
124720	1	575.7137	-216.20	2	124720.	61.7	0	181-191	R.VLIAAHGNLSR.G	
36490	1	775.1887	924.57	2	36490.0	36.3	1	181-195	R.VLIAAHGNLSRGIVK.H	
35941	1	597.2936	-30.26	2	35941.0	52.0	1	232-241	R.FLGDEETVRK.A	
37330	1	717.9048	747.56	3	37330.0	57.9	2	232-251	R.FLGDEETVRKAMEAVAAQGK.A	
35436	1	975.1800	-320.65	1	35436.0	67.1	0	242-251	K.AMEAVAAQGK.A	

Labeling quantitative proteomics

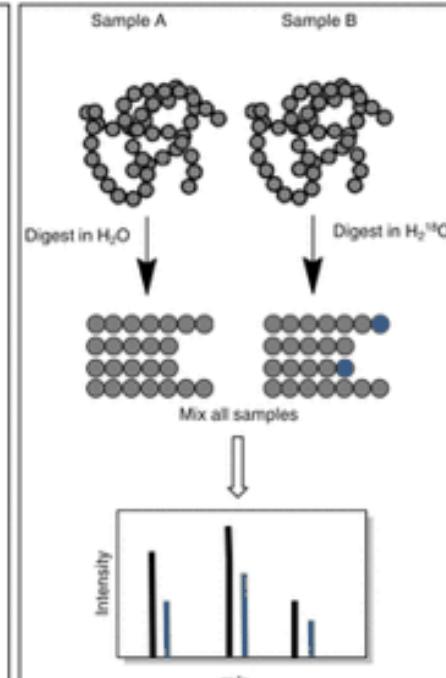
Chemical labeling



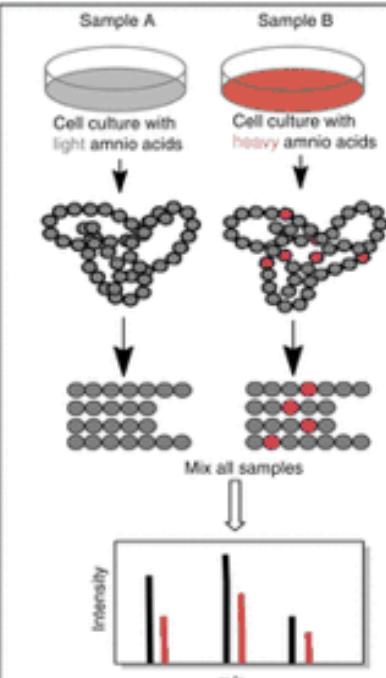
Isobaric tags labeling



Enzymatic labeling



Metabolic labeling

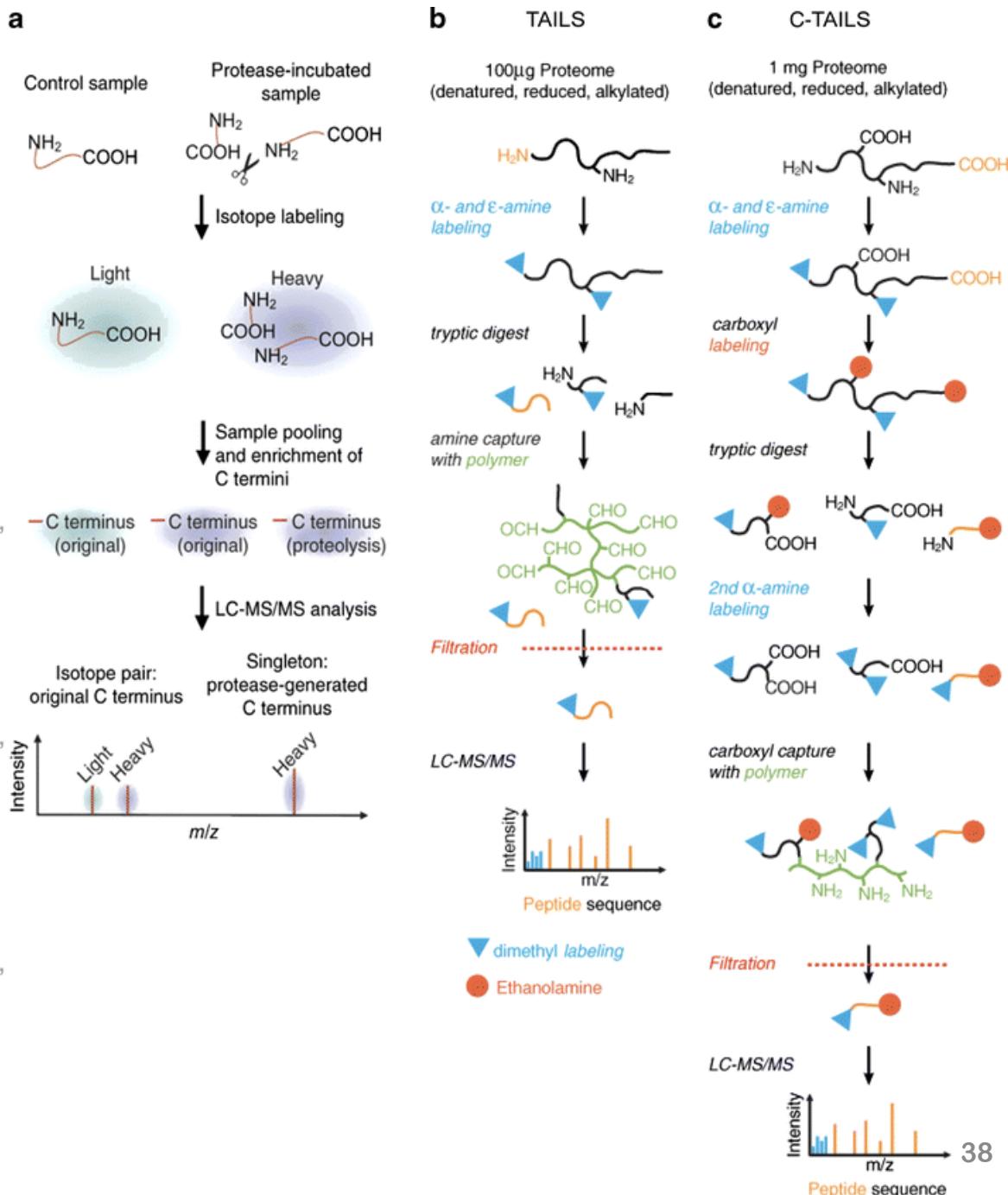
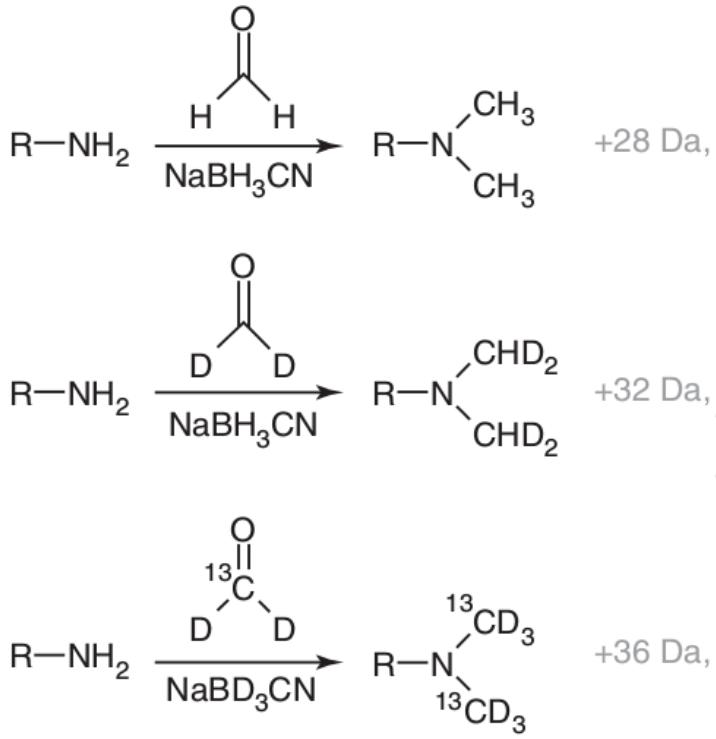


(e.g., dimethyl labeling)

(e.g., iTRAQ)

(e.g., SILAC)

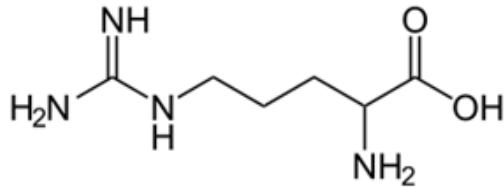
Dimethyl labeling of primary amines



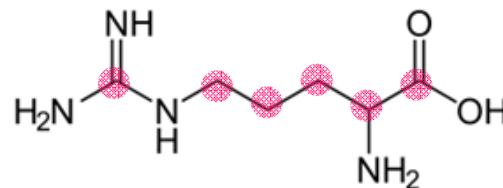
Stable Isotopes: Variant Amino Acids

● ^{13}C

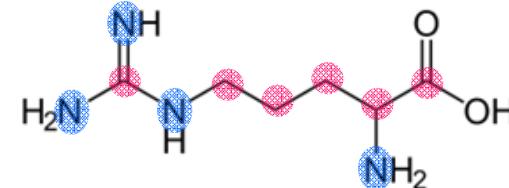
● ^{15}N



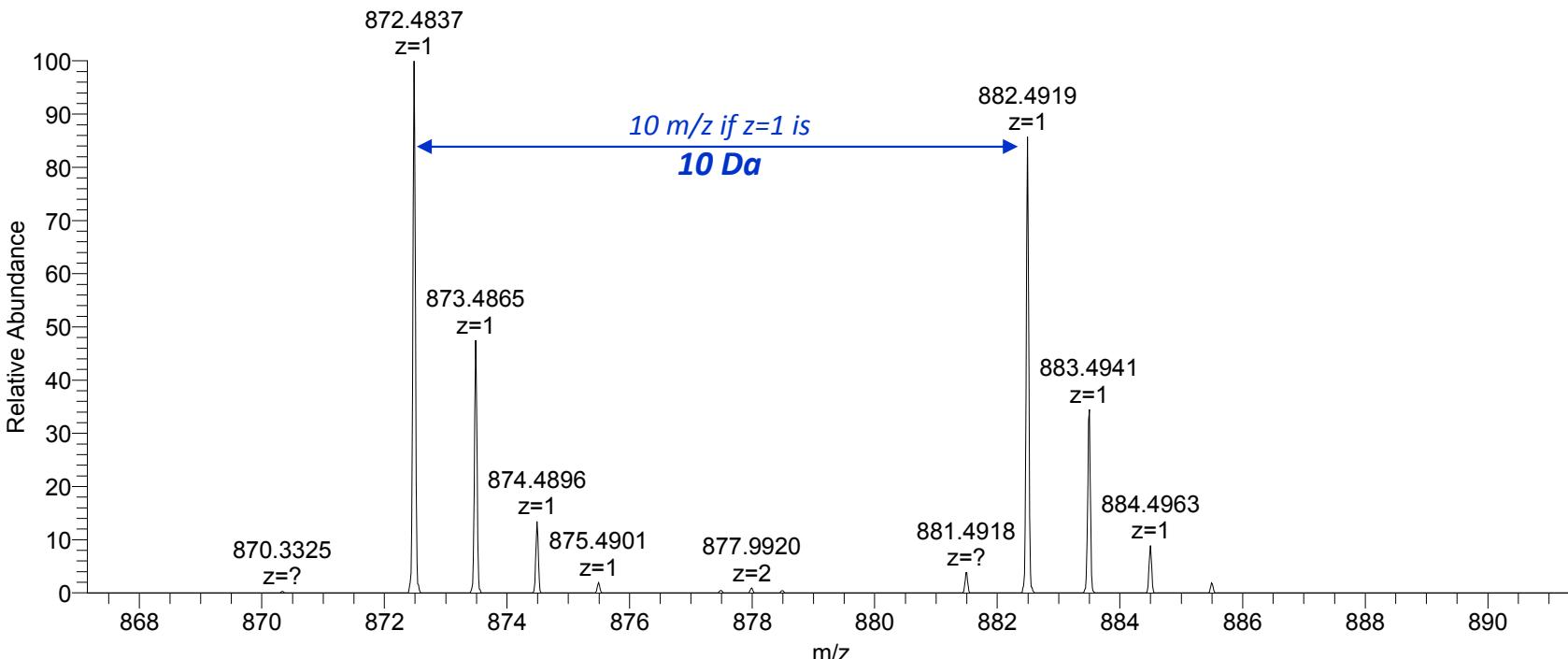
Arg₊₀ (Light)



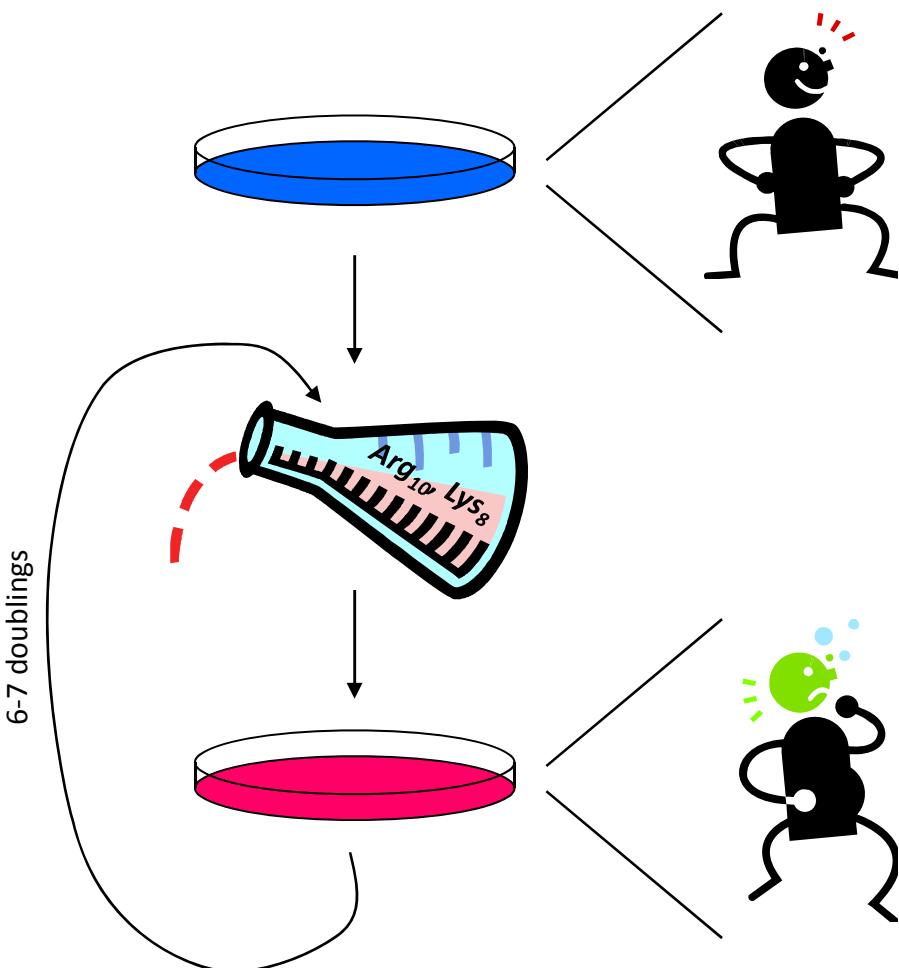
Arg₊₆ (Medium)



Arg₊₁₀ (Heavy)



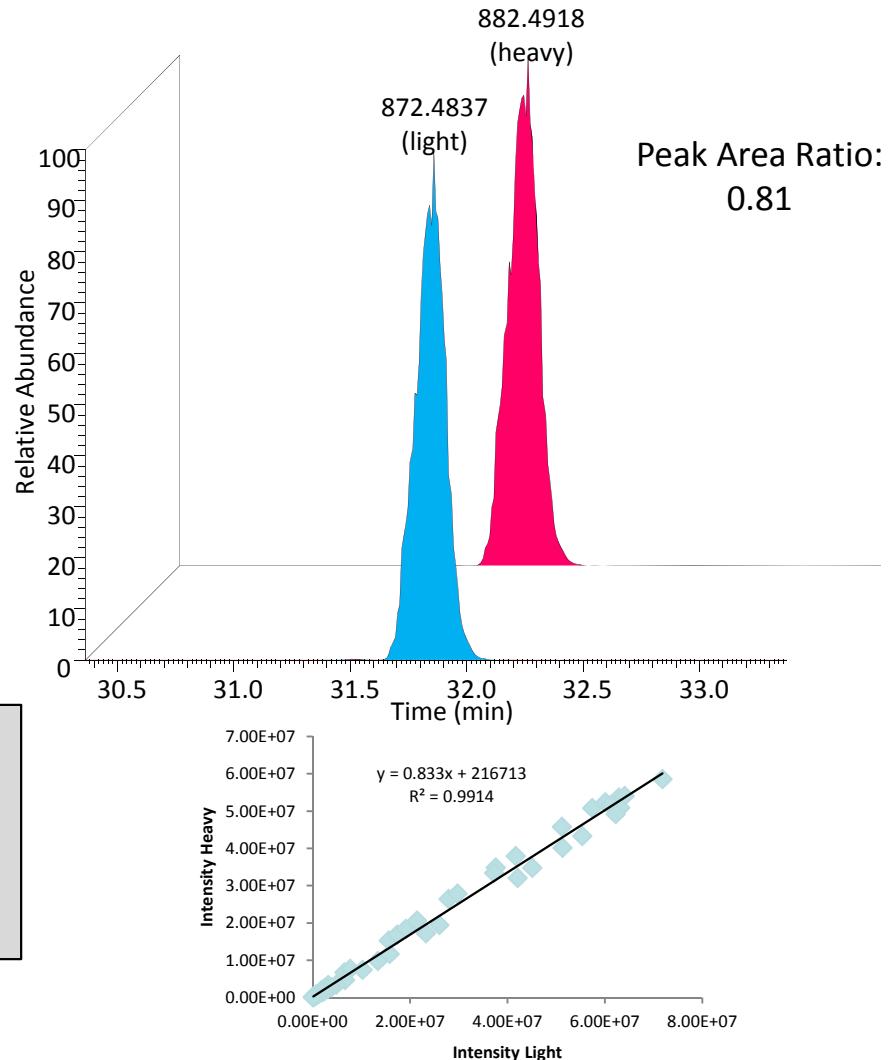
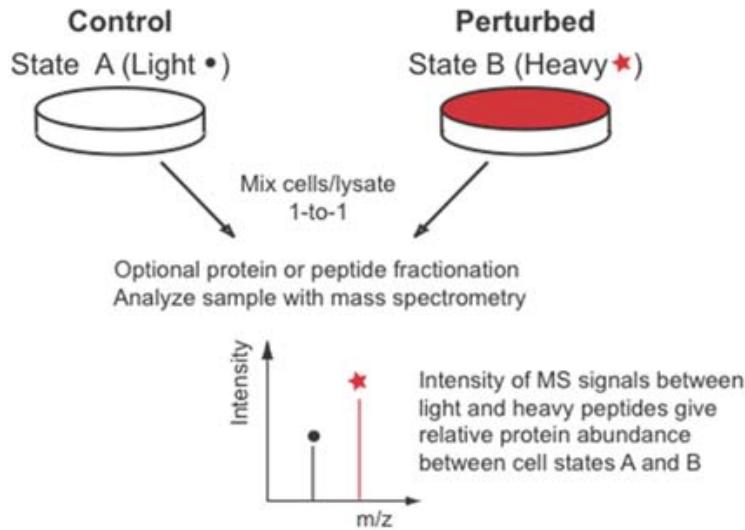
Stable Isotopes: Let cells do the work



- Feed cells growth medium with “heavy” amino acids
- Proteins incorporate these amino acids
- All proteins are slightly “heavier” but otherwise biochemically equivalent
- Choice of Lys and Arg typical because of trypsin

SILAC: Built-in Standardization!

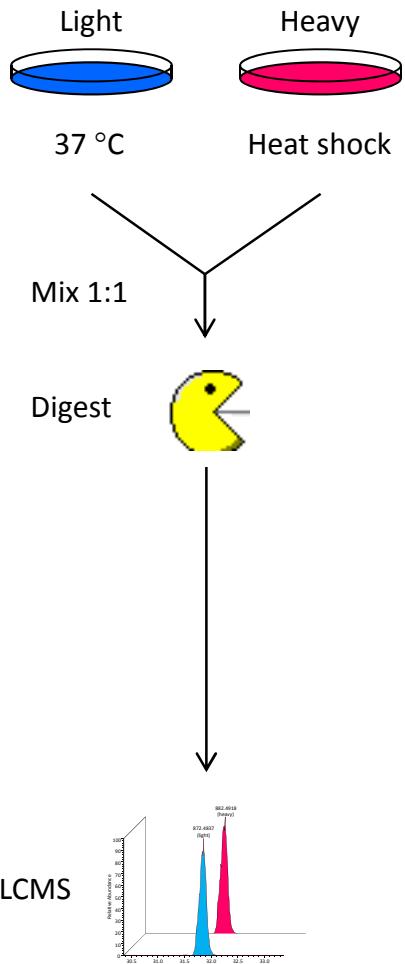
■ Stable Isotope Labeling of Amino acids in Culture



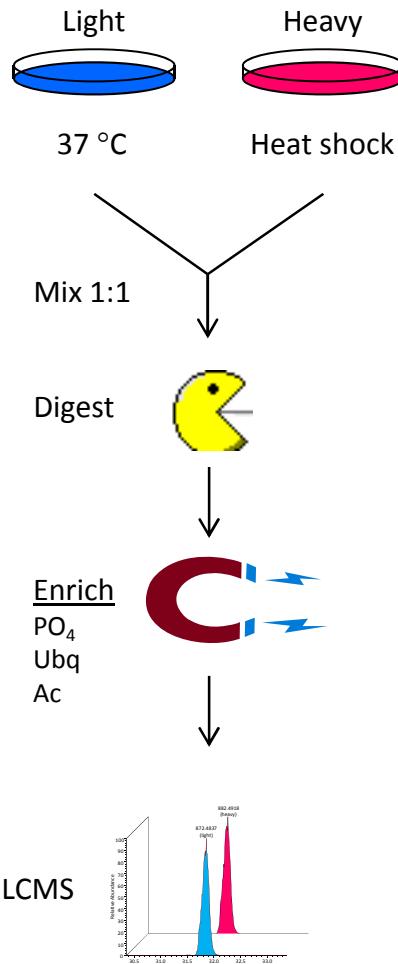
- Output is a *ratio* of a protein or a peptide in 2-3 conditions
 - Relative abundance

SILAC: Application Examples

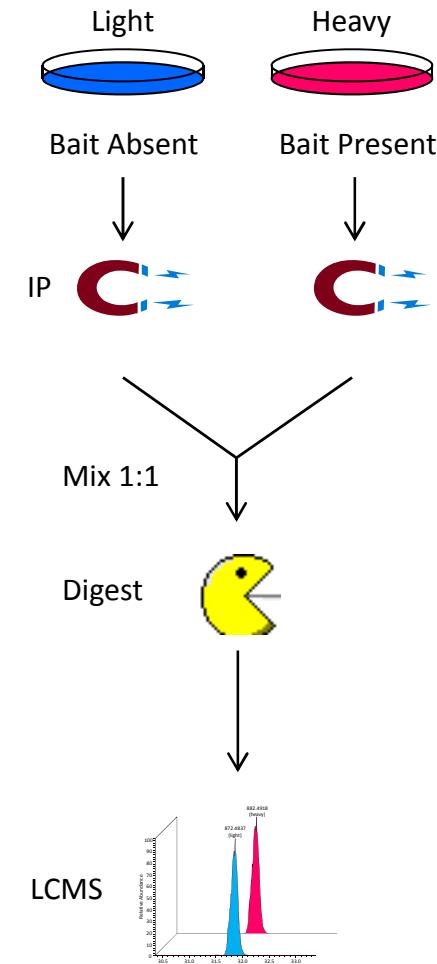
Protein Expression Profiling



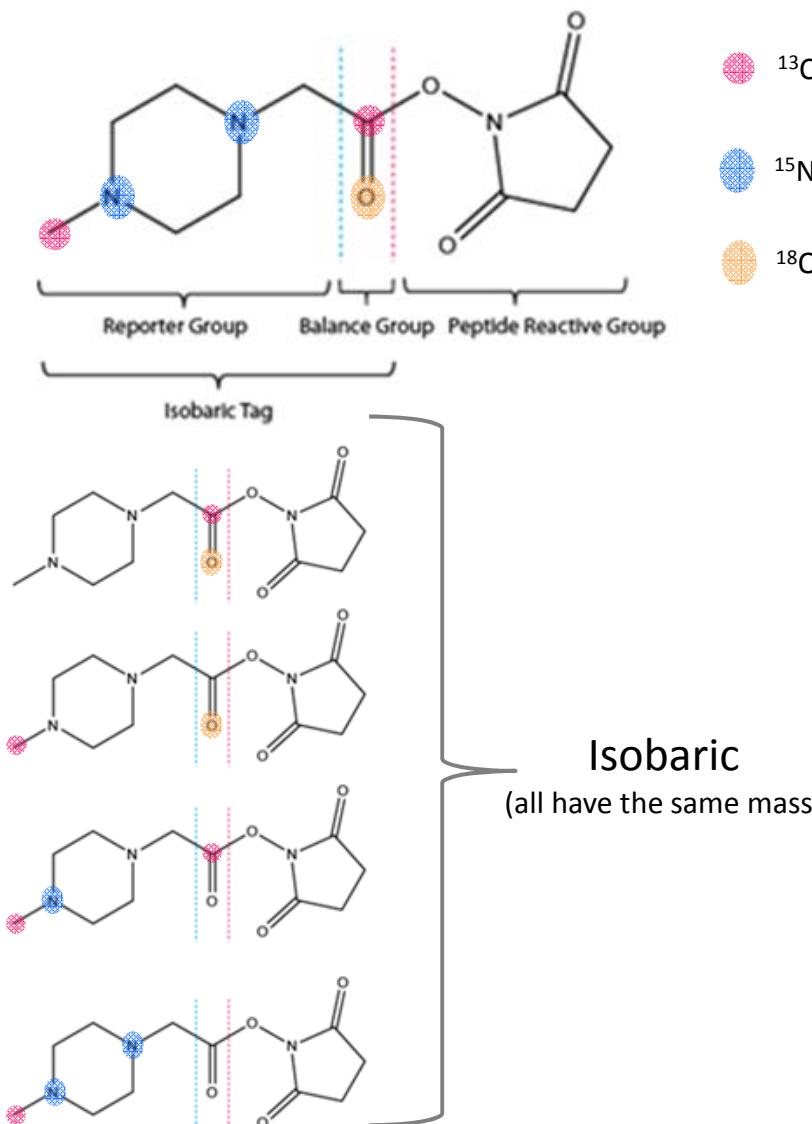
PTM Expression Profiling



Protein Interaction Profiling

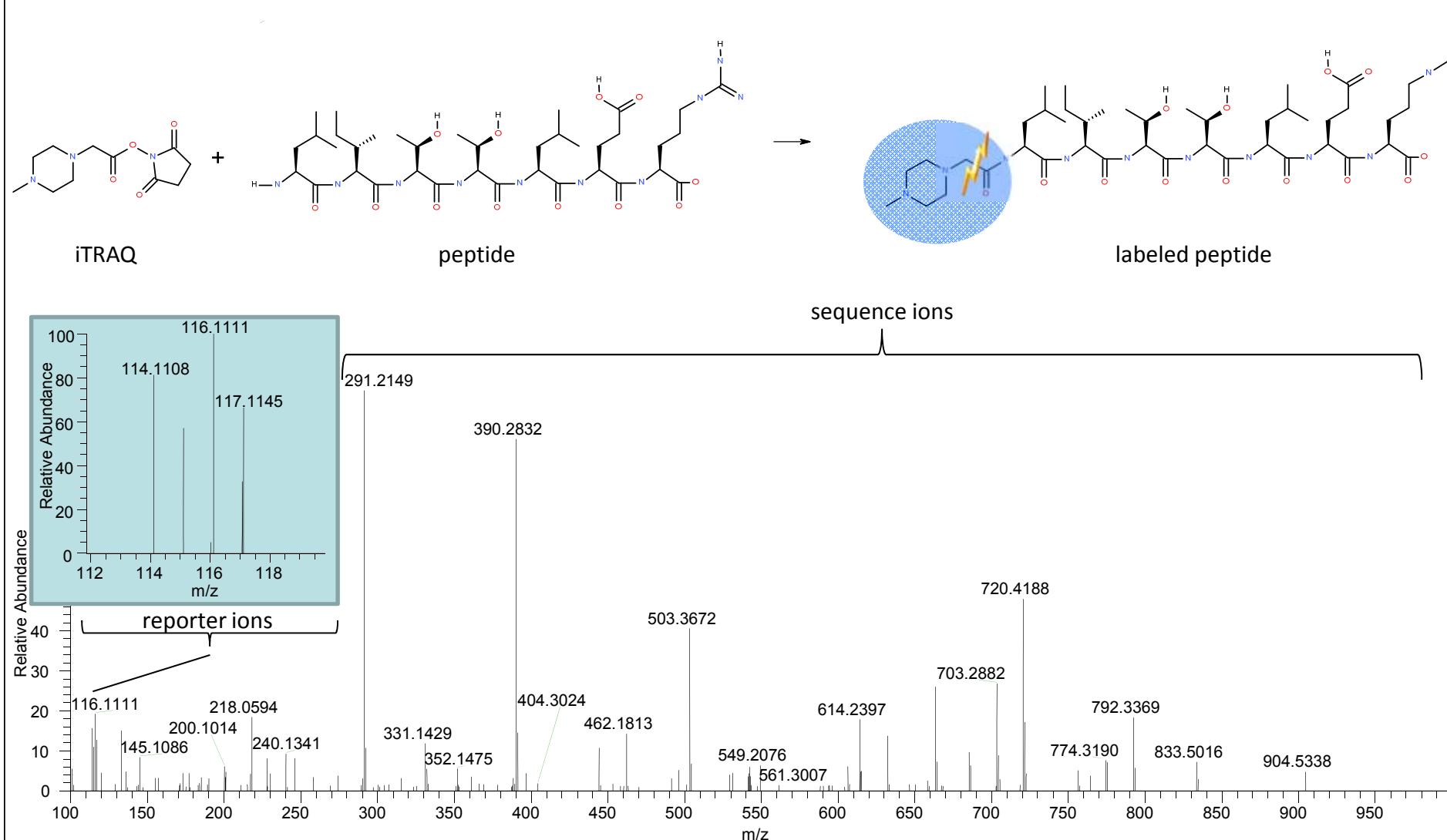


iTRAQ – Structures and Introduction

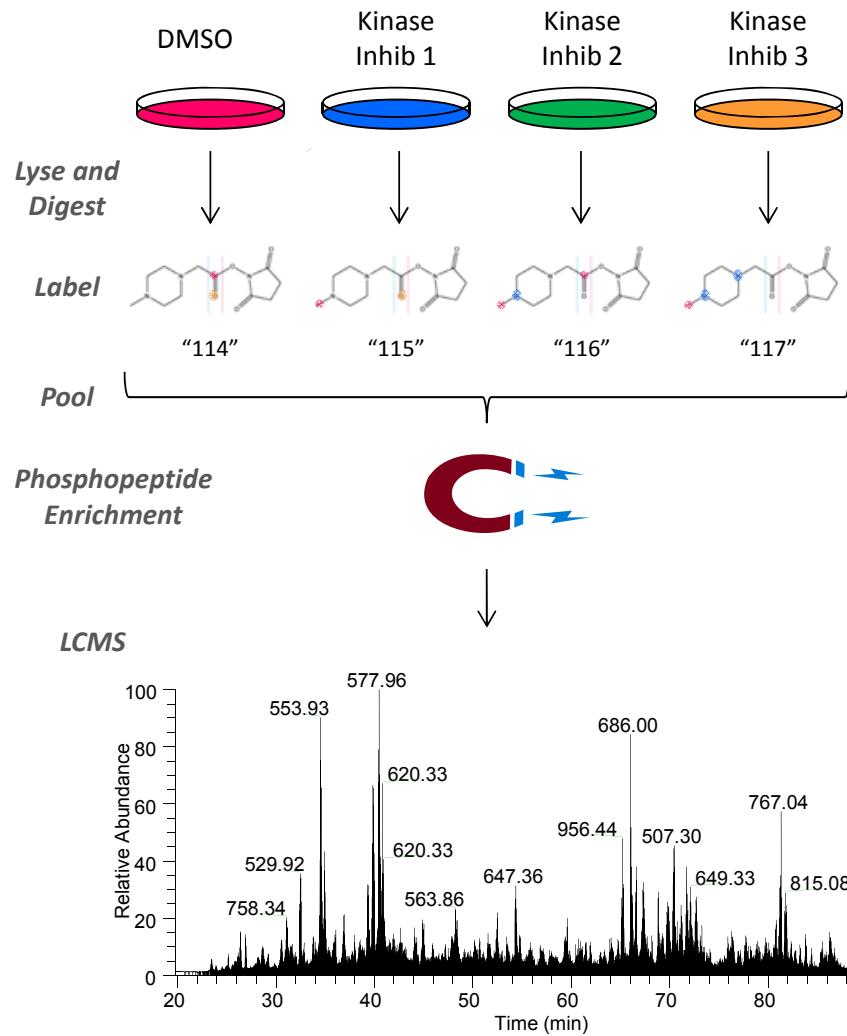


- iTRAQ- isobaric Tags for Relative and Absolute(?) Quantification

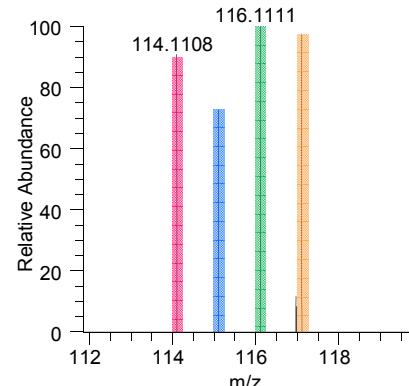
iTRAQ Fragmentation and Spectrum



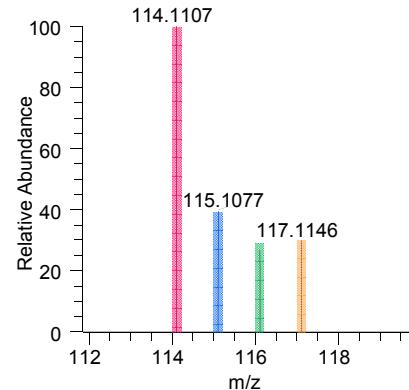
iTRAQ Experimental Example



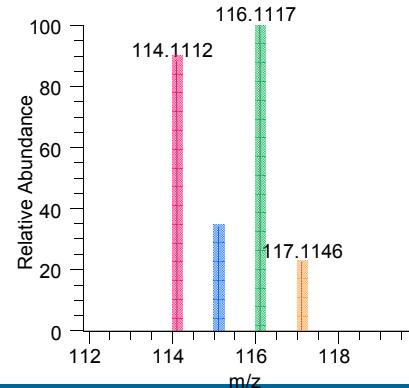
*Peptide #1:
No effect*



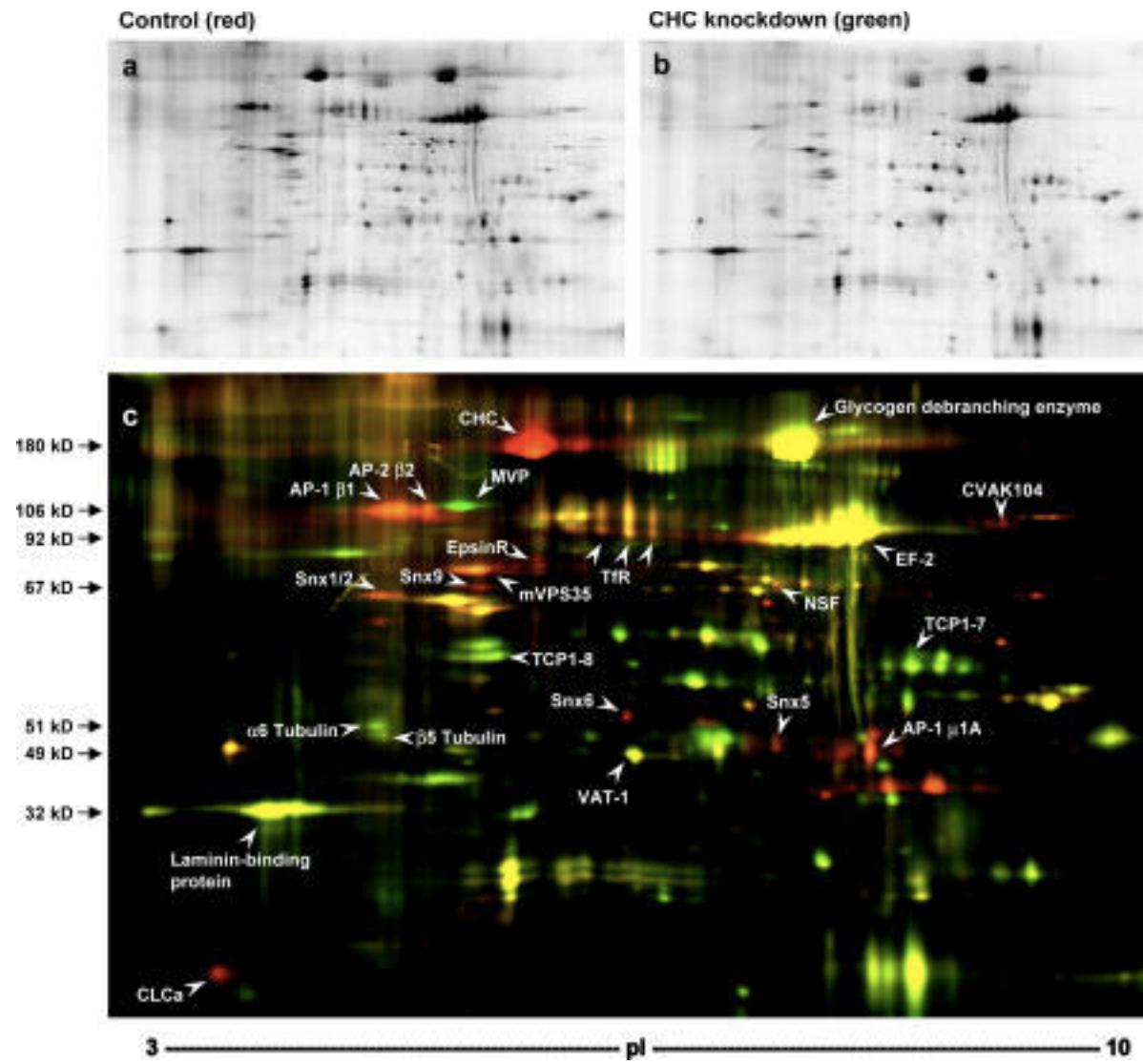
*Peptide #2:
Sensitive to
all inhibitors*



*Peptide #3:
Sensitive to
inhibitors 1 & 3*



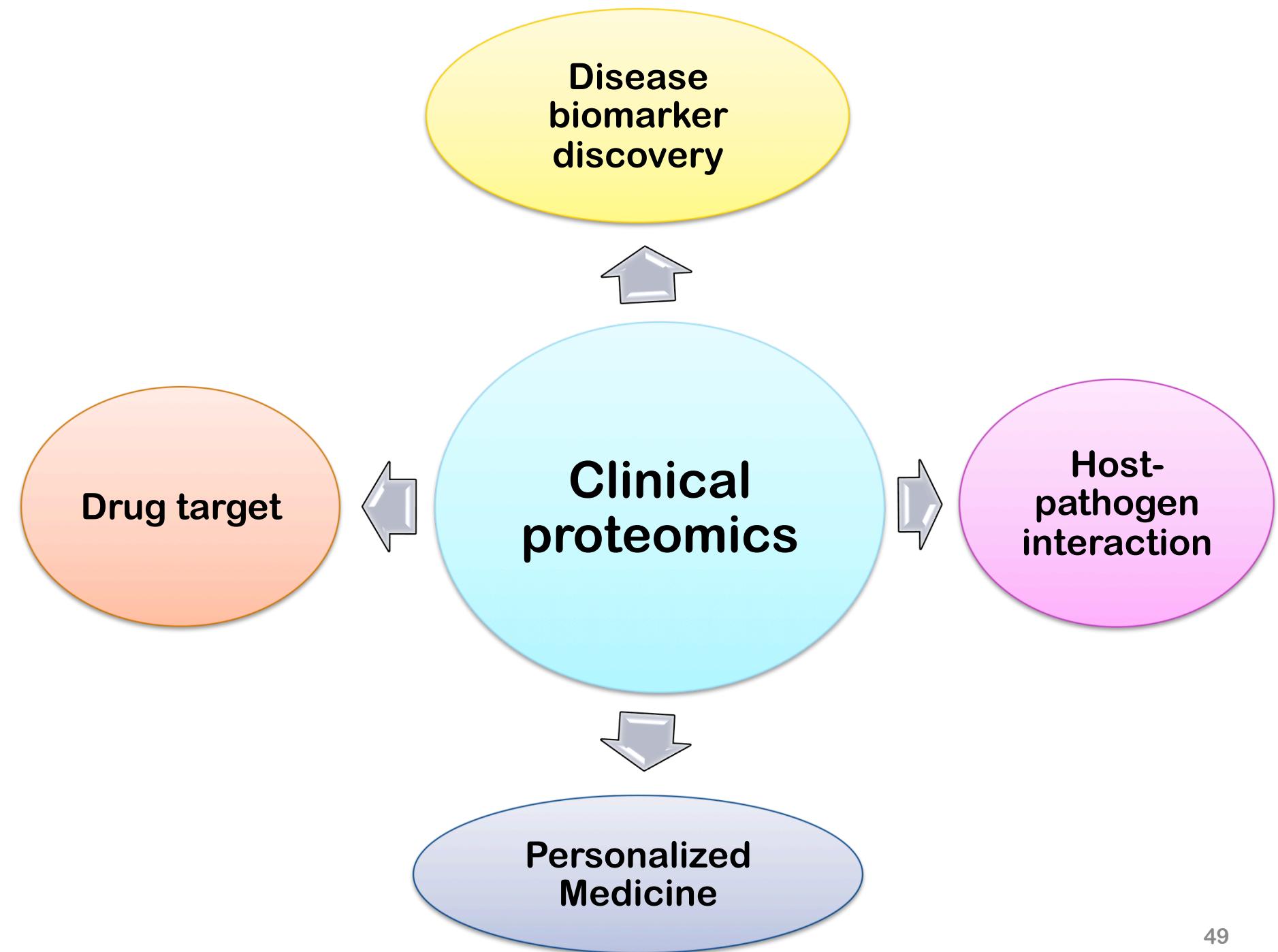
2D-DIGE (Difference gel electrophoresis)



Comparison of quantitative methods

Methods	Pros	Cons
Label-free	<ul style="list-style-type: none">- Easy biochemical workflows- High quantitative proteome coverage- No labeling reagents necessary- High dynamic range	<ul style="list-style-type: none">- Low accuracy- Moderate reproducibility- Robust sample preparation required- No multiplexing possible
SILAC	<ul style="list-style-type: none">- Multiplexing capacity- High accuracy- High proteome coverage- Good reproducibility	<ul style="list-style-type: none">- Limited to model systems- Low dynamic range- High resolution MS necessary- Additional costs for labeled amino acids
iTRAQ	<ul style="list-style-type: none">- High multiplexing capacity (up to 10-plex)- High accuracy- Well-established labeling kits- Reactivity towards different amino acids	<ul style="list-style-type: none">- Expensive labeling reagents- Additional labeling step- Low dynamic range- High resolution MS necessary
2D-DIGE	<ul style="list-style-type: none">- Detection of isoforms and PTMs- High accuracy and sensitivity- High dynamic range- Multiplexing (2-plex + internal std)	<ul style="list-style-type: none">- Expensive labeling dyes- Low quantitative proteome coverage- High sample amounts needed- High effort

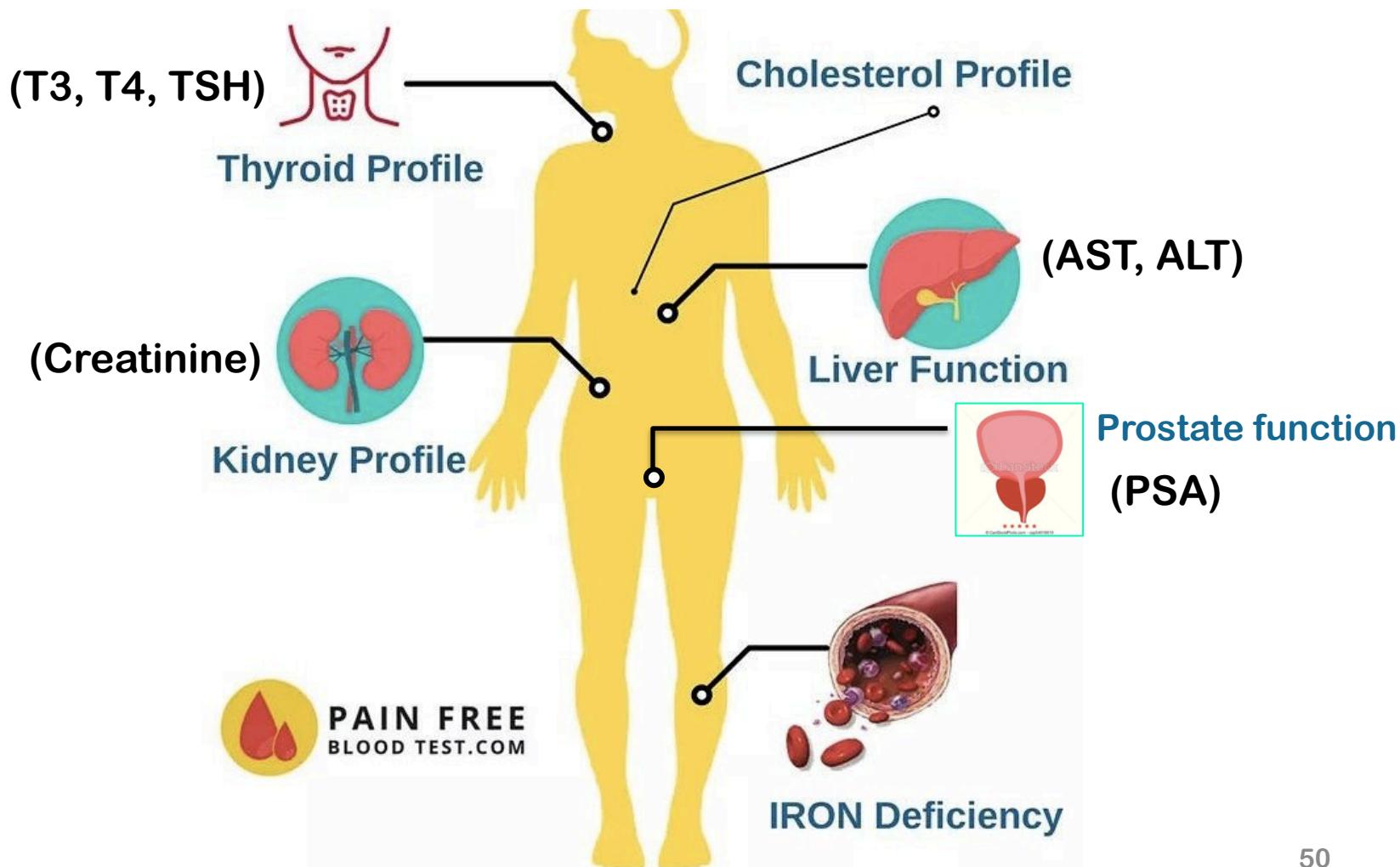
Applications of proteomics



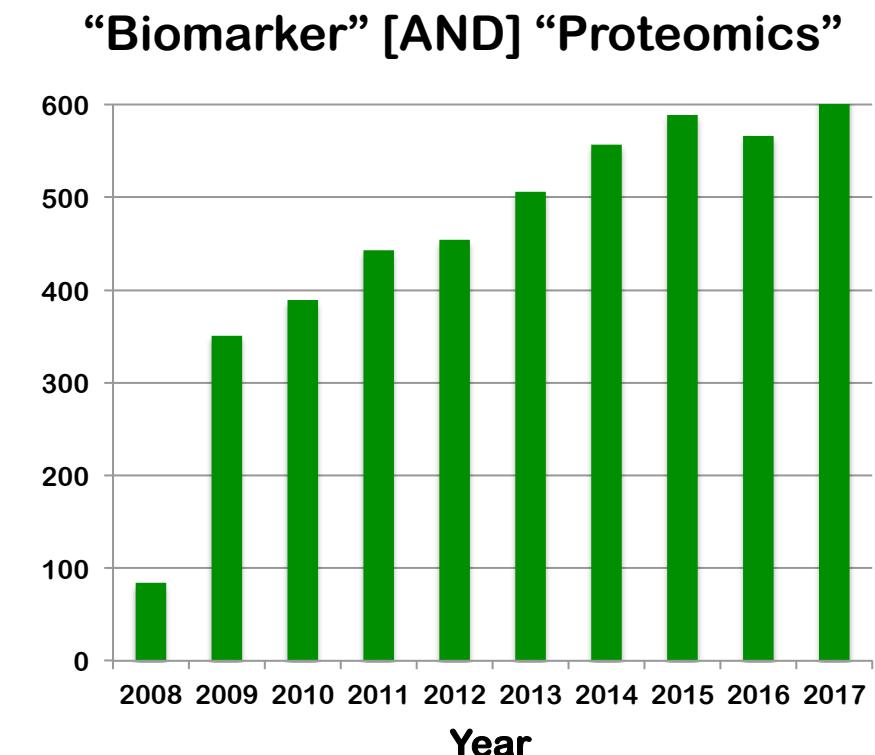
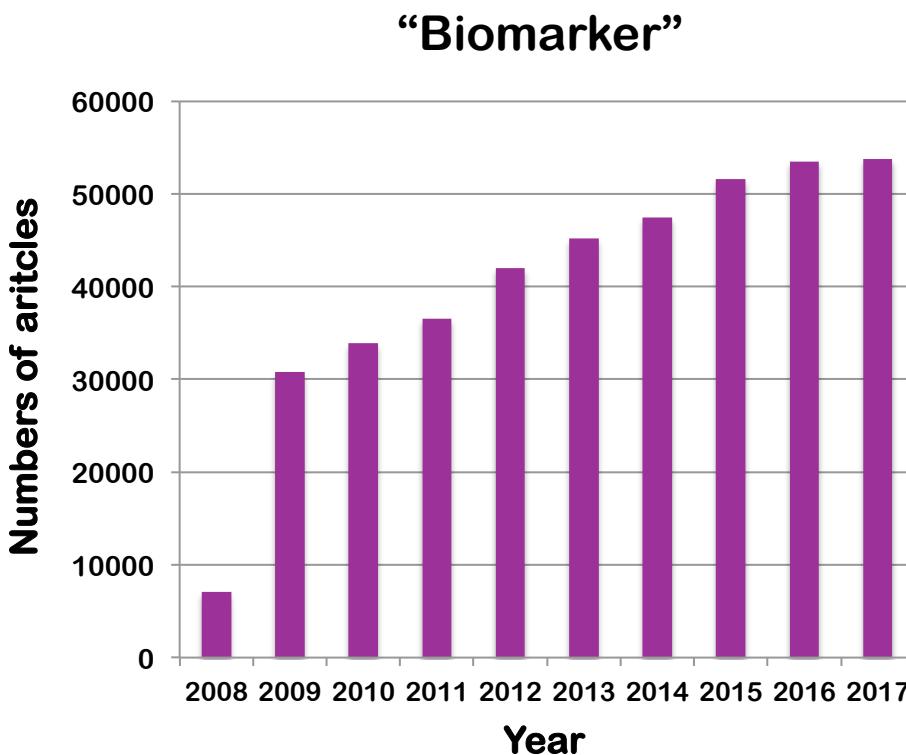
Biomarkers

- Biomarkers are defined as “**measurable characteristics that reflect physiological, pharmacological, or disease processes**”, according to the European Medicines Agency.

<http://genomemedicine.com/content/5/2/17>



Numbers of “Biomarker” and “Proteomics” publications in Pubmed



Many proteins proposed as biomarkers but very few introduced into clinical practice ($\leq 1/\text{year}$)

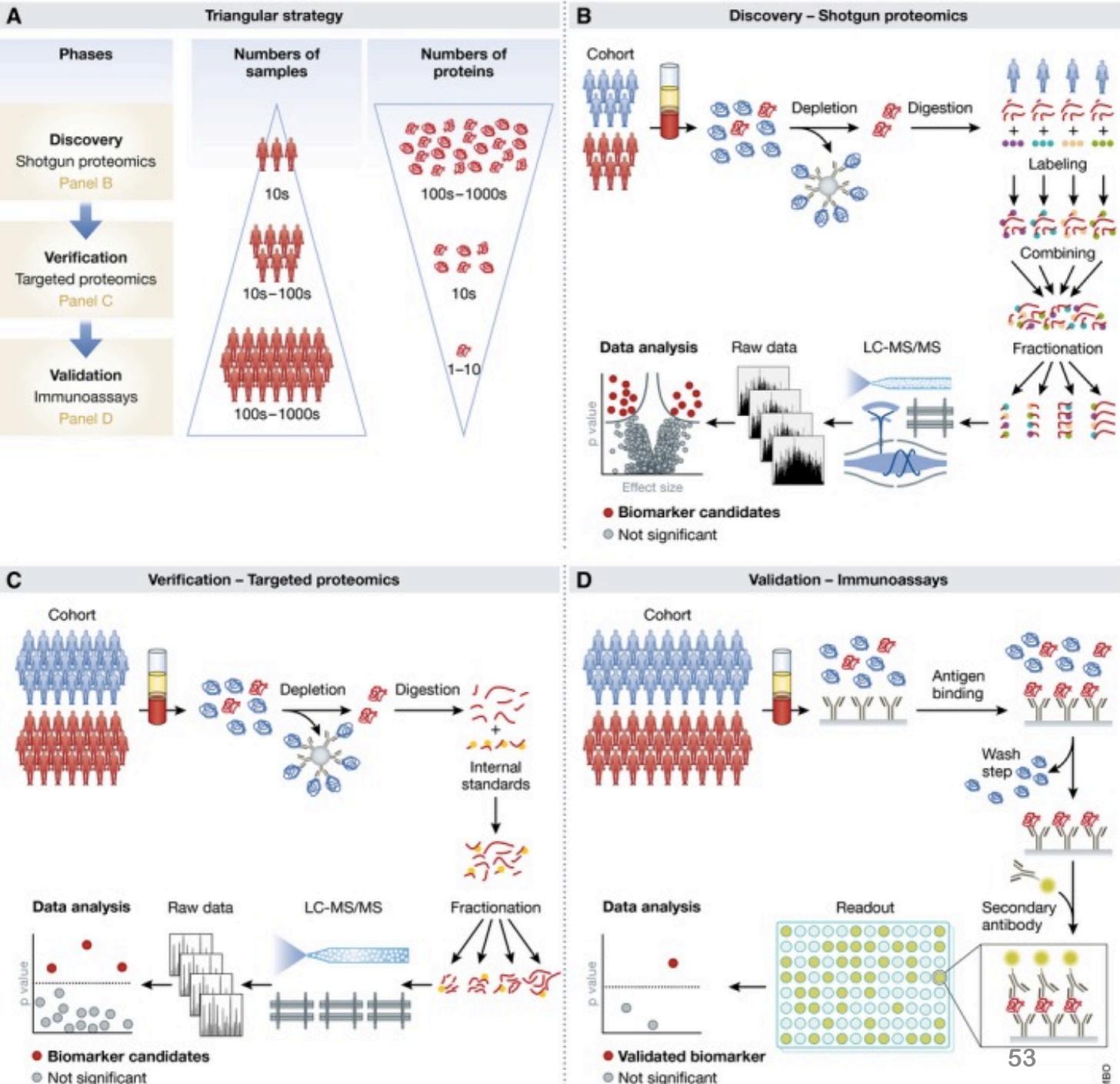
Why? Can we do better?

Obstacles to progress in proteomics-based biomarker development

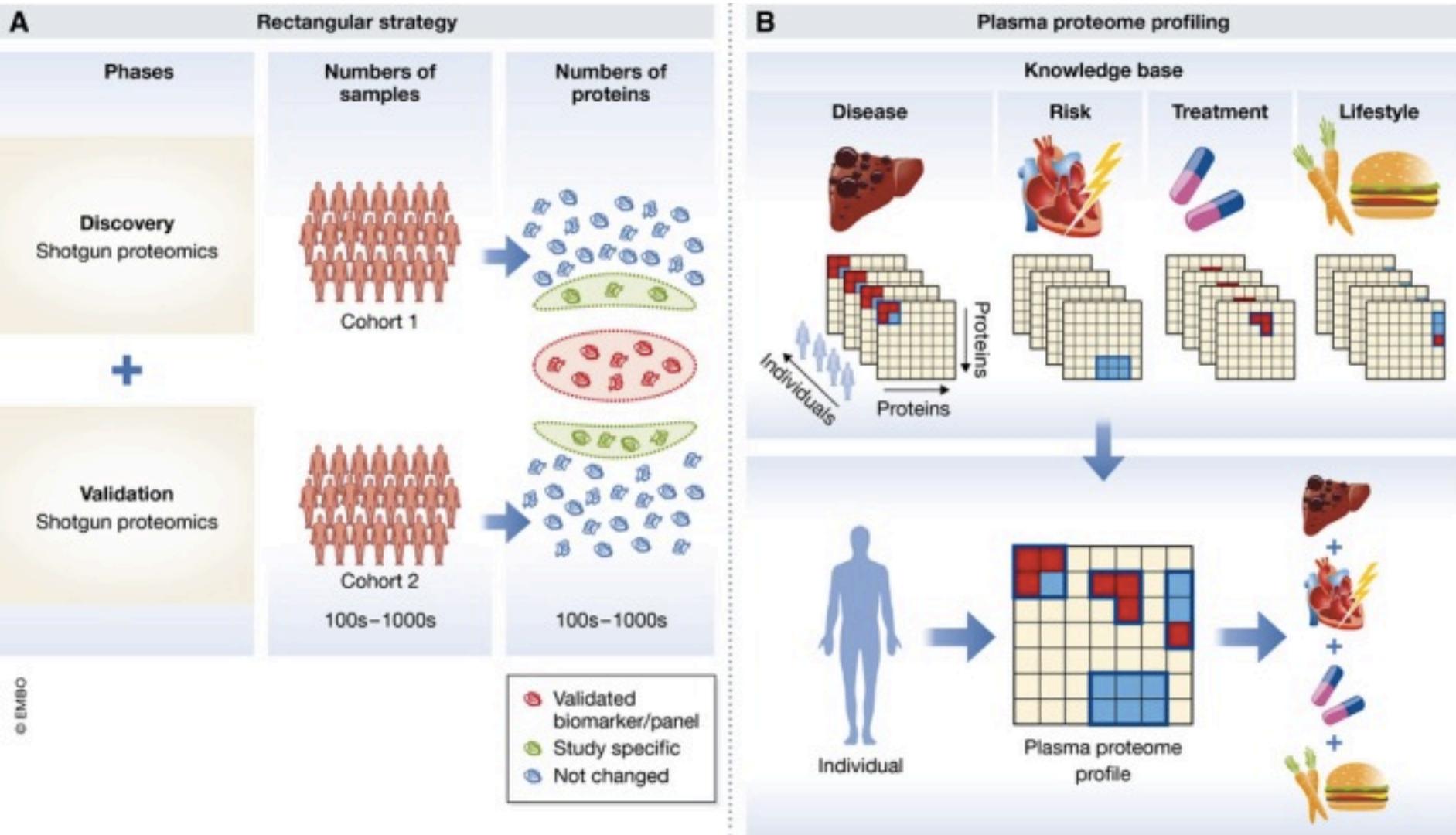
- **MS-platforms inadequate for the task**
 - Difficulty in repeatedly and precisely measuring large number of peptides/proteins over 10^8 concentration range
 - Low number of patient samples – high false-discovery rate (FDR)
- **Absence of coordinated teams, including biostatisticians, clinicians, proteomic specialists → Poor study design**
- **Multiple ad hoc, statistically indefensible data analysis methods**
- **Few expert proteomics labs willing/able to focus on clinical sample analysis**
 - Many studies have described readily detectable, abundant proteins with no specific disease association
- **Need for methods to quantify large numbers of peptides/proteins from discovery in 100's of patient samples**
 - Must be robust, quantitative, highly multiplexed, sensitive, specific

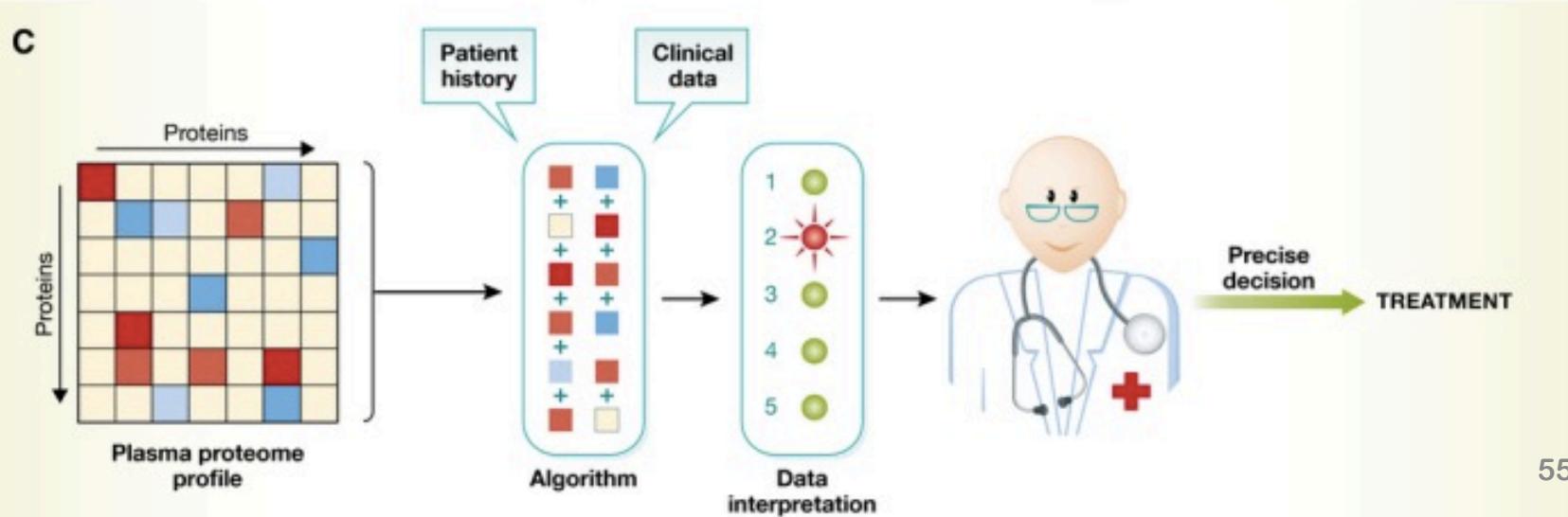
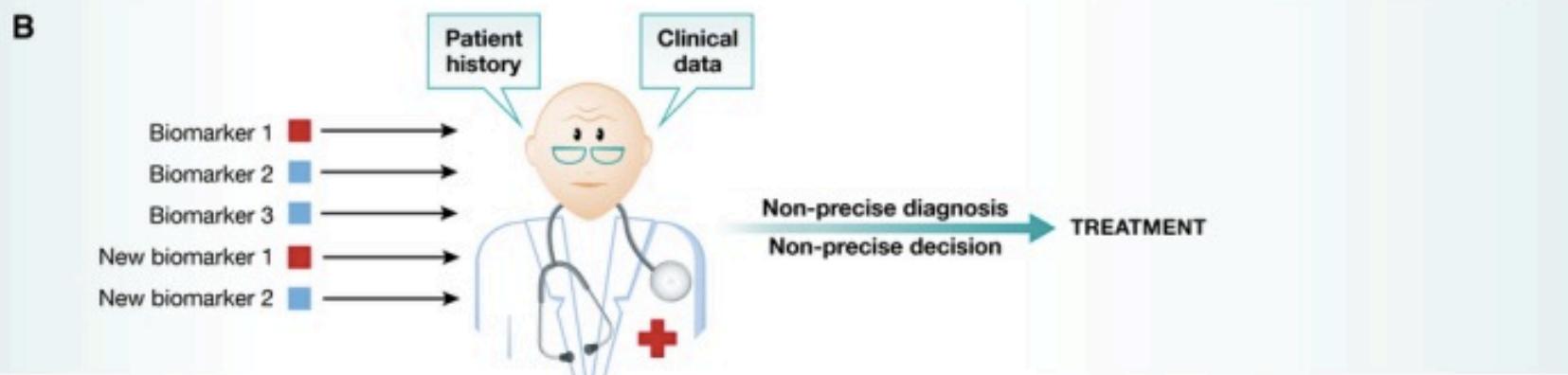
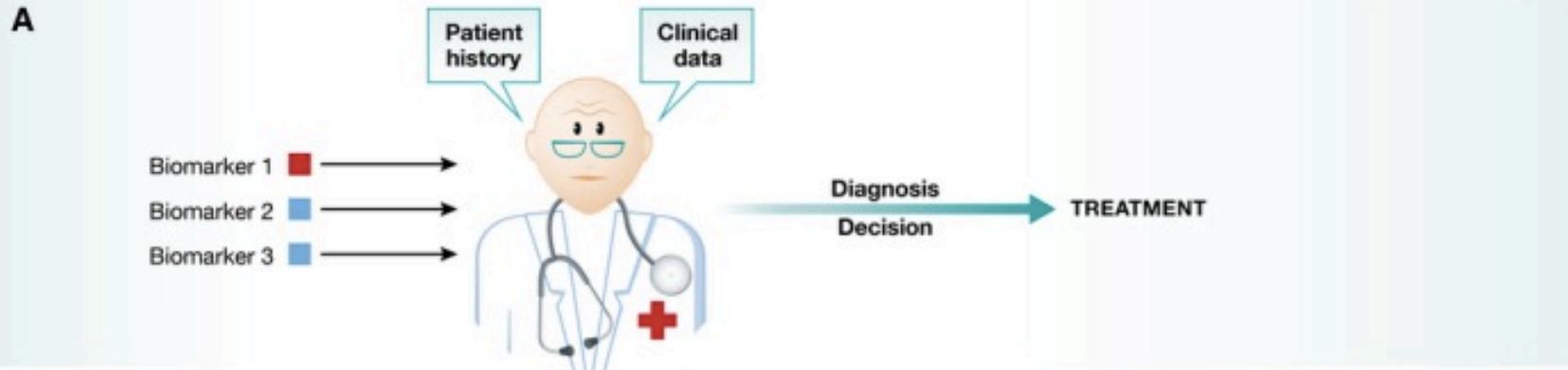
Current paradigms in plasma biomarker research

(Triangular workflow)



Future paradigms in plasma biomarker research with improved technology (Rectangular workflow)





A specific and very recent example of proteomics application



NATURE COMMUNICATIONS | (2018)9:3598 | DOI: 10.1038/s41467-018-05696-2 | www.nature.com/naturecommunications

ARTICLE

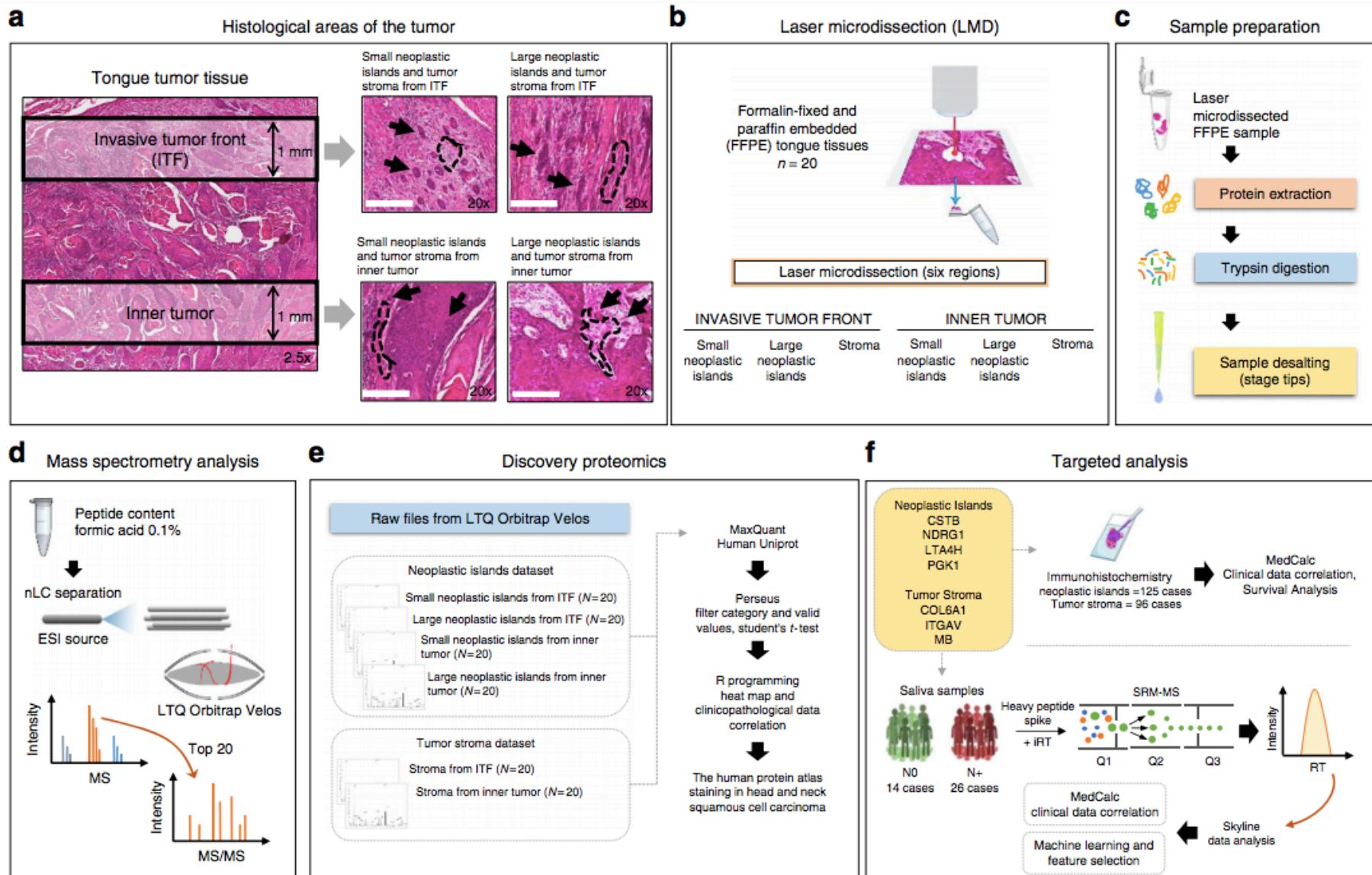
DOI: 10.1038/s41467-018-05696-2

OPEN

Combining discovery and targeted proteomics reveals a prognostic signature in oral cancer

Carolina Moretto Carnielli ¹, Carolina Carneiro Soares Macedo^{1,2}, Tatiane De Rossi¹, Daniela Campos Granato¹, César Rivera ^{1,2}, Romênia Ramos Domingues¹, Bianca Alves Pauletti¹, Sami Yokoo¹, Henry Heberle ³, Ariane Fidelis Busso-Lopes¹, Nilva Karla Cervigne^{2,4}, Iris Sawazaki-Calone⁵, Gabriela Vaz Meirelles¹, Fábio Albuquerque Marchi⁶, Guilherme Pimentel Telles⁷, Rosane Minghim³, Ana Carolina Prado Ribeiro^{8,9}, Thaís Bianca Brandão⁸, Gilberto de Castro Jr¹⁰, Wilfredo Alejandro González-Arriagada¹¹, Alexandre Gomes¹², Fabio Penteado¹², Alan Roger Santos-Silva², Márcio Ajudarte Lopes², Priscila Campioni Rodrigues^{13,14}, Elias Sundquist^{13,14}, Tuula Salo^{13,14,15}, Sabrina Daniela da Silva^{16,17}, Moulay A. Alaoui-Jamali¹⁷, Edgard Graner², Jay W. Fox¹⁸, Ricardo Della Coletta² & Adriana Franco Paes Leme ¹

DOI: 10.1038/s41467-018-05696-2



DOI: 10.1038/s41467-018-05696-2

Key references

- <https://www.broadinstitute.org/proteomics>
- Shen X., Shen S., Qu J. (2016) Labeling and Label-Free Shotgun Proteomics Quantification in the Research of Cardiovascular Diseases. In: Agnetti G., Lindsey M., Foster D. (eds) Manual of Cardiovascular Proteomics. Springer, Cham
- Patterson SD, Aebersold RH. 2003. Proteomics: the first decade and beyond. *Nat Genet* 33 Suppl:311-23.
- Geyer PE, Holdt LM, Teupser D, Mann M. 2017. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol* 13:942.
- Carnielli CM, Macedo CCS, De Rossi T, Granato DC, Rivera C, Domingues RR, Pauletti BA, Yokoo S, Heberle H, Busso-Lopes AF, Cervigne NK, Sawazaki-Calone I, Meirelles GV, Marchi FA, Telles GP, Minghim R, Ribeiro ACP, Brandao TB, de Castro G, Jr., Gonzalez-Arriagada WA, Gomes A, Penteado F, Santos-Silva AR, Lopes MA, Rodrigues PC, Sundquist E, Salo T, da Silva SD, Alaoui-Jamali MA, Graner E, Fox JW, Coletta RD, Paes Leme AF. 2018. Combining discovery and targeted proteomics reveals a prognostic signature in oral cancer. *Nat Commun* 9:3598.

Special thanks to...My MS teacher

Professor Kym Francis Faull, the director of Pasarow Mass Spectrometry Laboratory, University of California, Los Angeles



**David Geffen
School of Medicine**

Thank you for your attention



River view from SiMPC
10th floor, Srisavarindira Building